(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 April 2005 (07.04.2005)

PCT

(10) International Publication Number
WO 2005/030985 A2

(51) International Patent Classification[7]: C12Q 1/00

(21) International Application Number:
PCT/GB2004/004103

(22) International Filing Date:
24 September 2004 (24.09.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/505,970	25 September 2003 (25.09.2003)	US
0322493.8	25 September 2003 (25.09.2003)	GB

(71) Applicant *(for all designated States except US)*: **DEVGEN N.V.** [BE/BE]; Technologiepark 30, B- 9052 Gent-Zwijnaarde (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DE WILDE, Gert Jules Hector** [BE/BE]; Dr. Armand Rubbensstraat 25, B- 9240 Zele (BE). **SAUNDERS, Michael John Scott** [GB/BE]; 132 rue Berkendael, B- 1050 Brussels (BE).

(74) Agents: **BALDOCK, Sharon, Claire** et al.; BOULT WADE TENNANT, Verulam Gardens, 70 Gray's Inn Road, LONDON WC1X 8BT (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2005/030985 A2

(54) Title: USE OF AMINO ACID SEQUENCES INVOLVED IN THE ELONGATION OF FATTY ACIDS IN IDENTIFYING AND/OR DEVELOPING COMPOUNDS FOR PREVENTING AND/OR TREATING METABOLIC DISEASES

(57) Abstract: The present invention relates to methods for the identification and/or the development of compounds that can be used to prevent and/or to treat metabolic diseases, and to the use of amino acid sequences that are involved in the elongation of fatty acids in such methods. The invention also relates to compounds that have been identified and/or developed using said methods, to pharmaceutical compositions that contain such compounds, and to the use of said compounds in the preparation of pharmaceutical compositions, in particular of pharmaceutical compositions for the prevention and/or treatment of metabolic diseases. The invention further relates to compounds that can interact with amino acid sequences that are involved in the elongation of fatty acids, to pharmaceutical compositions that contain such compounds, and to the use of said compounds in the preparation of pharmaceutical compositions, in particular of pharmaceutical compositions for the prevention and/or treatment of metabolic diseases.

Use of amino acid sequences involved in the elongation of fatty acids in identifying and/or developing compounds for preventing and/or treating metabolic diseases.

The present invention relates to methods for the identification and/or the development of compounds that can be used to prevent and/or to treat metabolic diseases, and to the use of amino acid sequences that are involved in the elongation of fatty acids in such methods.

The invention also relates to compounds that have been identified and/or developed using said methods, to pharmaceutical compositions that contain such compounds, and to the use of said compounds in the preparation of pharmaceutical compositions, in particular of pharmaceutical compositions for the prevention and/or treatment of metabolic diseases.

The invention further relates to compounds that can interact with amino acid sequences that are involved in the elongation of fatty acids, to pharmaceutical compositions that contain such compounds, and to the use of said compounds in the preparation of pharmaceutical compositions, in particular of pharmaceutical compositions for the prevention and/or treatment of metabolic diseases.

Other aspects, embodiments, applications and advantages of the present invention will become clear from the further description below.

The present invention was established on the basis of the finding that amino acid sequences that are involved in the elongation of fatty acids (as further described below) can be used as (potential) "target(s)" for *in vitro* and/or *in vivo* interaction with chemical compounds and other factors (with the term "*target*" having its usual meaning in the art, vide for example the definition given in WO 98/06737). Consequently, compounds and/or factors that have been identified as interacting with amino acid sequences that are involved in the elongation of fatty acids (e.g. using the methods described herein) may be useful as active agents in the pharmaceutical field, and in particular for the prevention and/or treatment of metabolic diseases.

Pathways for the *de novo* biosynthesis of fatty acids, as well as the amino acid sequences/enzymes that are involved in such pathways, have been described in the art. See for example:

- Leonard et al., Biochem. J. (2000), 350, p. 765-770), for the elongation pathway in *Saccharomyces cerevisiae;*
- Tvrdik et al., The Journal of Cell Biology, Vol.149 (2000), p.707-717; Tvrdik et al., J. Biol. Chem., Vol. 272, No.50, p. 31738-31746 (1997); Moon et al., J. Biol. Chem., Vol. 276, No. 48, p. 45358-45366 (2001); Matsuzaka et al., J. Lipid. Res., Vol. 43 (2002), 911-920; and Moon and Horton, J. Biol. Chem., Vol. 278, No.9, 6335-7343 (2003), for the elongation pathway in mice and humans.

In animals, fatty acids are generally synthesized *de novo* from acetyl-CoA and stearyl-CoA through a series of reactions mediated by acetyl-CoA carboxylase (also referred to as "*ACC*") and fatty acid synthase (also referred to as "*FAS*"). Reference is *inter alia* made to Moon et al, *supra*.

In mammals, this *de novo* synthesis principally results in, and ends with (i.e. for about 90%, with about 10% of higher fatty acids such as stearic acid being formed), the fatty acid palmitic acid (which, in accordance with standard fatty acid nomenclature, is often designated as "16:0", in which "16" denotes the number of carbon atoms in the fatty acid carbon chain - i.e. C_{16} - and "0" denotes the number of unsaturated bonds, i.e. 0), see again Moon et al., *supra*. This is released into the cytosol, from where it is either further elongated - i.e. to stearic acid (18:0, which is the main saturated product resulting from the elongation reaction) and/or is desaturated (e.g. from 16:0 to 16:1 for palmitic acid, or from 18:0 to 18:1 for stearic acid, etc.).

The elongation of fatty acids is performed in the mitochondria and more commonly in the endoplasmatic reticulum by a complex of multiple enzymes, and involves a cycle of condensation (referred to herein as the "First Step"), reduction (referred to herein as the "Second Step"), dehydration (referred to herein as the "Third Step") and reduction (referred to herein as the "Fourth Step"), which are as follows (see Moon and Horton, *supra*):

First Step: condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA;

Second Step: reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA;

Third Step: dehydration of 3-hydroxyacyl-CoA to *trans*-2,3-enoyl-CoA;

Fourth Step: reduction of *trans*-2,3-enoyl-CoA to saturated acyl-CoA.

This cycle results in the addition of two carbon atoms to the fatty acid chain, compared to the starting fatty acid. The fatty acid chain thus elongated may then either be further elongated through another cycle of condensation, reduction, dehydration and reduction; and/or may be desaturated. In turn, the further elongated and/or desaturated fatty acids thus obtained may then again be further elongated and/or desaturated. In this way, starting from palmitic acid as the main product from the *de novo* synthesis by the *FAS/ACC* system, a whole range of saturated and unsaturated fatty acids may be synthesized.

However, some families of biologically important fatty acids cannot be synthesized by animals, because they lack some of the required enzymes. For instance, when it comes to the desaturation of saturated fatty acids and/or the further desaturation of unsaturated fatty acids, animals lack the desaturases required for introducing a double bond beyond the 9^{th} position in the carbon chain, and can only insert additional double bonds between an existing double bond and the terminal carboxylgroup.

Thus, for instance, animals cannot synthesize (the precursors for) the fatty acids of the so-called "(n-6) series" (in which, in accordance with standard biochemical nomenclature, "*n*" denotes the number of carbon atoms and "*6*" denotes the position of the last double bond, counting from the methyl group determining the metabolic family) and the "(n-3) series", which include for example linoleic acid (18:2(n-6)), the primary precursor for arachidonic acid (20:4 (n-6)) and other fatty acids of the (n-6) series, and *alpha*-linolenic acid (18:3(n-3)), the main precursor for fatty acids of the (n-3) series. Animals must therefore take up these essential fatty acids from plant sources as part of their diet (mainly in the form of linoleic acid and alpha-linolenic acid); the fatty acids

thus taken up may then be elongated and/or further desaturated to provide other fatty acids of the (n-6) and (n-3) series, again via the steps outlined above.

The cycle of elongation steps described above is performed by a system of multiple enzymes comprising:

First Step: a condensing enzyme;

Second Step: a β (beta)-oxoacyl-CoA reductase;

Third Step: a β (beta)-hydroxyacyl-CoA dehydrase;

Fourth step: a trans-2-enoyl-CoA reductase;

which enzymes are also collectively referred to as "*elongation enzymes*" or "*elongases*" (see for example Moon et al, *supra*), and sometimes also as the "*fatty acid chain elongation system*" or "*FACES*" (see for example Leonard et al, *supra*).

It is generally assumed that the condensation (the First Step above) is catalysed by different members from a group of essentially similar but distinct condensation enzymes, each of which has substrate specificity for a fatty acid (or for a series of fatty acids) with a specific number (or range) of carbon atoms in the fatty acid chain. For example, in *S. cerevisiae*, it has been shown that there is one condensation enzyme - called "*ELO1*"- which has specificity for the elongation of C14:0 fatty acids into C16:0 fatty acids (compared to for instance the elongation of C16:0 into C18:0, see Toke et al., J. Biol. Chem., Vol.271, No. 31, p. 18413-18422 (1996)) and that there are two other condensation enzymes - called "*ELO2*"/"*FEN1*" and "*ELO3*"/"*SUR4*", respectively - which have specificity for the condensation/elongation of fatty acids up to C_{24} and for the condensation/elongation of C_{24} fatty acids into C_{26} fatty acids, respectively (see Oh et al., J. Biol. Chem., Vol.272, No. 28, p. 17376-17384 (1997).

The functional equivalents of ELO2 and ELO3 from mice have been described as "*Cig30*" and "*Ssc1*", respectively (see Tvrdik et al. (2000), *supra*). Also, a condensation enzyme from mice with specificity for the conversion of C_{16} to C_{18} fatty acids - called

"*LCE*" - and the sequence of its human ortholog have been identified (see Moon et al., *supra*, in particular the alignment in Figure 2 on page 45361).

The art also describes several assays for determining the activity and/or (substrate) specificity of elongases with respect to different fatty acids, for example the "*in vitro fatty acid elongation assay*" described in Moon et al. (*supra*), the "*fatty acid elongation assay*" described by Tvrdik et al., (2000, *supra*) and/or the "*fatty acid elongation assay*" described by Matsuzaka et al. (*supra*). As indicated therein and as further described below, such assays may also be performed with for example microsomal preparations, which may contain two or more or essentially all enzymes involved in the First Step to Fourth Step above, so as to measure the activity of the entire pathway, or the influence of a single enzyme on said pathway.

The (proposed) biological function of mammalian elongation enzymes can also be determined/confirmed by (over)expression of cDNA encoding said enzymes in S. *cerevisiae* (as performed for the human elongase HELO1 (= ELOVL5) by Leonard et al., *supra*) or in a mutant of *S. cerevisiae* that lacks one or more of the native elongations enzymes ELO1, ELO2 and/or ELO3 (see Toke et al. and Oh et al., both *supra*).

The nucleotide sequences and proposed amino acid sequences for the human condensation enzymes involved in the First Step above - which are also collectively referred to as the "*ELOVL*" family - have also been described in the art, i.e. as follows:

- ELOVL6 , which is the human ortholog of the mouse elongase LCE, which has specificity for the conversion of C_{14} and C_{16} saturated and mono-unsaturated fatty acids, such as for the conversion of C_{16} to C_{18} fatty acids, and in particular for the conversion of palmitic acid (16:0) to stearic acid (18:0). The nucleotide sequence of ELOVL6 is given in SEQ ID NO: 1 (mRNA sequence) and SEQ ID NO: 2 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO: 3. See also Genkbank accession numbers NM_024090.1 (gi 13129087) for the nucleotide sequence and NP_076995.1 (gi 13129088) for the amino acid sequence. Reference is also made to Moon et al., *supra.*; By (mRNA sequence) it is meant that the nucleotide sequences listed are derived from the full mRNA sequence (including introns). By (mRNA-coding sequence) it is meant that the nucleotide sequences listed are derived

from the mRNA sequence which encodes the amino acid sequence, that is with introns removed.

- ELOVL3, which is the human ortholog of the mouse elongase *Cig30* and the *S. cerevisiae* elongase ELO2, with specificity for very long chain fatty acids (up to C_{26}). The nucleotide sequence of ELOVL3 is given in SEQ ID NO: 6 (mRNA sequence) and SEQ ID NO: 7 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO:8. See also Genkbank accession numbers NM_152310 (gi 23097309); nucleotide sequence) and NP_689523.1 (gi 23097310; amino acid sequence). Reference is also made to Tvrdik (1997), *supra*, and to Moon et al., *supra*.
- ELOVL5 (also called "HELO 1"), which is involved in the elongation of polyunsaturated fatty acids (PUFA's). The nucleotide sequence of ELOVL5 is given in SEQ ID NO: 9 (mRNA sequence) and SEQ ID NO: 10 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO:11. See also Genkbank accession numbers NM_021814 (gi 21361903) for the nucleotide sequence and NP_068586.1 (gi 11464975) for the amino acid sequence. amino acid sequence). Reference is also made to Leonard et al., *supra.*;
- ELOVL4, which is involved in the elongation of very long chain fatty acids. The nucleotide sequence of ELOVL5 is given in SEQ ID NO: 12 (mRNA sequence) and SEQ ID NO: 13 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO:14. See also Genkbank accession numbers NM_022726 (gi 21362099) for the nucleotide sequence and NP_073563.1 (gi 12232379) for the amino acid sequence.
- ELOVL2, which is the human ortholog of the mouse elongase *Ssc2*, which has been shown to accept arachidonic acid and eicosapentaneoic acid as its substrates. The nucleotide sequence of ELOVL2 is given in SEQ ID NO: 15 (mRNA sequence) and SEQ ID NO: 16 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO:17. See also Genkbank accession numbers NM_17770 (gi 8923311) for the nucleotide sequence and NP_060240.1 (gi 8923312) for the amino acid sequence. Reference is also made to Tvirdik (2000), *supra*, and to Moon et al., *supra*.
- ELOVL1, which is the human ortholog of the mouse elongase *Ssc1* and the *S. cerevisiae* elongase *ELO 3*, with specificity for very long chain fatty acids. The

nucleotide sequence of ELOVL1 is given in SEQ ID NO: 18 (mRNA sequence) and SEQ ID NO: 19 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO:20. See also Genkbank accession numbers NM_22821 (gi 13489092) for the nucleotide sequence and NP_073732.1 (gi 13489093) for the amino acid sequence. Reference is also made to Tvirdik (2000), *supra*, and to Moon et al., *supra*.

In general, all the condensation enzymes described above contain a histidine-rich motif (H*XX*HH), and also contain an ELO-domain (as determined by SMART™ analysis (see http://smart.embl-heidelberg.de/) and/or PFAM protein search™ (see for example http://pfam.wustl.edu/). For example, the ELO domain in the amino acid sequence of SEQ ID NO: 3 comprises amino acids 10-265.

In addition to the six enzymes from the ELOVL family above, bio-informatic analysis has also identified a further human amino acid sequence – shown in SEQ ID NO: 5 - that contains both an H*XX*HH motif (amino acids 24-28) as well as an "ELO" domain on PFAM protein search™ (amino acids 1-94). Although the invention is not limited to any specific hypothesis or explanation, it is assumed that this amino acid sequence is also involved in the elongation of fatty acids, and in particular in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA (i.e. the First Step above). The nucleotide sequence for this hypothetical protein (partial CDS) is given in SEQ ID NO: 4 and the amino acid sequence is given in SEQ ID NO: 5. Reference is also made to Genbank accession number AL137506 (gi: 6808153) for the nucleotide sequence, the CAB 70777.1 (gi: 6808154) for the amino acid sequence. These sequences and their use in the methods described herein form further aspects of the invention.

Compared to the condensation enzymes involved in the First Step, it is assumed that the subsequent steps of reduction (the Second Step above), dehydration (the Third Step above) and reduction (the Fourth Step above) are performed by enzymes that have less substrate specificity than the condensation enzymes described above.

The human reductases that are involved in the Second Step and Fourth Step are described by Moon and Horton, *supra*:

- "*3-ketoacyl-CoA reductase*" or "*KAR*", which is involved in the reduction of 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA (the Second Step above).

The nucleotide sequence of KAR is given in SEQ ID NO: 21 (mRNA sequence) and SEQ ID NO: 22 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO: 23. See also Genkbank accession numbers NM_016142 (gi 7705854) for the nucleotide sequence and NP_057226.1 (gi 7705855) for the amino acid sequence. KAR has also been described in the art under the name "*17-beta-hydroxysteroid dehydrogenase*" of "*HSD17B12*";

- "*trans-2,3-enoyl-CoA reductase*" or "TER", which is involved in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA (the Fourth Step above). The nucleotide sequence of TER is given in SEQ ID NO: 24 (mRNA sequence) and SEQ ID NO: 25 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO: 26. See also Genkbank accession numbers NM_004868 (gi 4759061) for the nucleotide sequence and NP_004859.1 (gi 4759062) for the amino acid sequence. TER has also been described in the art under the names "*synaptic glycoprotein 2*" or "*GPSN2*", and "*SC2*".

KAR (and its homolog NM_031463 of SEQ ID NO. 33, which has been found by bio-informatic analysis, see the Examples below) both contain an "*adh_short*" domain, as determined by PFAM analysis (amino acid residues 49-280 for KAR and amino acid residues 66-306 for NM_031463). TER contains a "*steroid_short*" domain, again as determined by PFAM analysis. (amino acid residues 12-157).

For the sake of completeness, it s mentioned herein that the desaturation of fatty acids (both saturated and unsaturated) is performed by enzymes which are referred to as "*fatty acid desaturases*". These *inter alia* include, but are not limited to, the enzymes known as the "*stearoyl-CoA-desaturases*" or "*SCD 's*". Examples of SCD's are SCD-1 and SCD-5 from humans, which are *inter alia* described in the International applications WO 01/66758 and WO 02/26944, respectively. Reference is also made to Ntambi et al., Curr. Opin. Lipifol 14:255-261 (2003) and the further references therein. However, such desaturases fall outside the scope of the present invention.

It has now been found that amino acids involved in the elongation of fatty acids can be used as targets (as described herein) in the field of metabolic disease.

In paricular, it has been found that amino acid sequences that are involved in one of the four steps of the elongation cycle of fatty acids – i.e. the First Step, Second Step,

Third Step and Fourth Step as described above - can be used as targets in metabolic disease, as further described hereinbelow. Throughout the specification "amino acid sequence" is defined to include any peptide, polypeptide, protein or enzyme. Also included within the definition are fragments of any of the above which retain functionality in terms of elongation of fatty acids.

According to one specific, but non-limiting embodiment of the invention, the amino acid sequence that is used in the invention is derived from multicellular animals, in particular from vertebrate animals, more in particular mammals, and even more in particular from humans, as further described hereinbelow.

According to one specific, but non-limiting aspect of the invention, the amino acid sequence that is used in the invention is chosen from the elongases, as described above, and more in particular in the elongases that are involved in the First Step, Second Step and/or Third Step above.

According to one specific, but non-limiting aspect of the invention, the amino acid sequence used in the invention is chosen from elongases involved in the First Step above, i.e. from the condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA.

More in particular, according to this aspect of the invention, the amino acid sequence used in the invention is chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for fatty acids with a length of the carbon chain of 24 carbon atoms or less (such as palmitic acid, stearic acid, linoleic acid and/or linolenic acid), compared to fatty acids with a length of the carbon chain of more than 24 carbon atoms.

Even more in particular, according to this aspect of the invention, the amino acid sequence used in the invention is chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for fatty acids with a length of the carbon chain of 20 carbon atoms or less (such as palmitic acid, stearic acid, linoleic acid and/or linolenic acid), compared to fatty acids with a length of the carbon chain of more than 20 carbon atoms.

Still more in particular, according to this aspect of the invention, the amino acid sequence used in the invention is chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for fatty acids with a length of the carbon chain of 18 carbon atoms or less, compared to fatty acids with a length of the carbon chain of more than 18 carbon atoms; and specifically for fatty acids with a length of the carbon chain of 16 or 18 carbon atoms, compared to fatty acids with a length of the carbon chain of more than 18 carbon atoms

Even still more in particular, according to this aspect of the invention, the amino acid sequence used in the invention is chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for fatty acids with a length of the carbon chain of 16 carbon atoms, compared to fatty acids with a length of the carbon chain of more than 16 carbon atoms (such as stearic acid). Some preferred, but non-limiting examples of such enzymes from mammals are LCE from mice (see Moon et al., *supra*) and its human equivalent (see again Moon et al., *supra*, as well as SEQ ID NOS: 1, 2 and 3).

Herein, the "substrate specificity" or "specificity" of an elongase can be determined using a suitable assay in a manner known per se, as will be clear to the skilled person. Generally, this will involve comparing two relevant substrates in such an assay, and determining which substrate is better accepted (e.g. more easily and/or faster converted) by the enzyme involved. For example, in the case of a condensation enzyme involved in the First Step above, the substrates may be tested in one of the *in vivo* or *in vitro* assays mentioned above and/or hereinbelow.

Generally, using such a suitable assay, the substrate specificity can be determined/expressed by means of the "turnover" of the substrate(s) involved (i.e. the amount of substrate converted by the enzyme within a given period of time, usually measured at a pre-determined concentration), with a higher turnover of a first substrate compared to a second substrate being an indication that the first substrate is better accepted by the enzyme involved than the second substrate. Alternatively, the substrate specificity can be expressed using the well-known K_m value, with a lower K_m value of a

first substrate compared a second substrate being an indication that the first substrate is better accepted by the enzyme involved than the second substrate.

Thus, according to one particular non-limiting aspect, an enzyme is considered to have "specificity" for a first substrate compared to a second substrate if, in a suitable assay, either (1) the turnover of the first substrate is at least 1.1 times, preferably at least 1.5 times, more preferably at least 2.5 times, even more preferably at least 5 times, and in particular 10 times or more the turnover for the second substrate (as measured in the same assay using essentially the same conditions; and/or (2) the K_m value of the first substrate is less than 90%, preferably less than 50%, more preferably less than 25%, even more preferably less than 10% of the K_m value for the second substrate (as measured in the same assay using essentially the same conditions).

According to one non-limiting embodiment of the aspect of the invention, the use of an amino acid sequence that is involved in - e.g. has specificity for - the elongation of unsaturated fatty acids that cannot be synthesized *de novo* by animals, or that are derived from fatty acid precursors that cannot be synthesized *de novo* by animals (i.e. by elongation and/or further desaturation), is less preferred.

In particular, according to this non-limiting embodiment of the aspect of the invention, the use of amino acid sequence(s) that are involved in the elongation of unsaturated fatty acids of the (n-3) family or (n-6) family are less preferred.

More in particular, the use of amino acid sequence(s) that are involved in the elongation of unsaturated fatty acids of the (n-3) family or (n-6) family with a length of the carbon chain of less than 20 carbon atoms (such as linoleic acid and alpha-linolenic acid) are less preferred.

Thus, according to one non-limiting embodiment of this aspect of the invention, the amino acid sequence(s) used as targets in the invention is/are chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for saturated fatty acids compared to unsaturated fatty acids of the (n-3) family and/or (n-6) family.

According to one particularly preferred, but non-limiting embodiment of this aspect of the invention, the amino acid sequence used in the invention is chosen from

condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for saturated fatty acids with 24 carbon atoms or less in the fatty acid chain (such as palmitic acid and stearic acid), compared to both (1) saturated and unsaturated fatty acids with more than 24 carbon atoms in the fatty acid chain; as well as (2) unsaturated fatty acids of the (n-3) family and/or (n-6) family with 20 carbon atoms or less in the fatty acid chain (such as linoleic acid and alpha-linolenic acid).

According to one particularly preferred, but non-limiting embodiment of this aspect of the invention, the amino acid sequence used in the invention is chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for saturated fatty acids with 20 carbon atoms or less in the fatty acid chain (such as palmitic acid and stearic acid), compared to both (1) saturated and unsaturated fatty acids with more than 20 carbon atoms in the fatty acid chain; as well as (2) unsaturated fatty acids of the (n-3) family and/or (n-6) family with 20 carbon atoms or less in the fatty acid chain (such as linoleic acid and alpha-linolenic acid).

According to one particularly preferred, but non-limiting embodiment of this aspect of the invention, the amino acid sequence used in the invention is chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for saturated fatty acids with 18 carbon atoms or less in the fatty acid chain, compared to both (1) saturated and unsaturated fatty acids with more than 18 carbon atoms in the fatty acid chain; as well as (2) unsaturated fatty acids of the (n-3) family and/or (n-6) family with 18 carbon atoms or less in the fatty acid chain.

More specifically, according to this preferred, but non-limiting embodiment, the amino acid sequence used in the invention is chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for palmitic acid and/or for stearic acid, compared to both (1) saturated and unsaturated fatty acids with more than 18 carbon atoms in the fatty acid chain; as well as (2) linoleic acid and/or alpha-linolenic acid.

Even more specifically, according to this preferred, but non-limiting embodiment, the amino acid sequence used in the invention is chosen from condensation enzymes involved in the First Step above that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for palmitic acid, compared to both (1) stearic acid; as well as (2) linoleic acid and/or alpha-linolenic acid. A preferred, but non-limiting examples of such enzymes from mammals are LCE from mice (see Moon et al., *supra*) and its human equivalent (see again Moon et al., *supra*, as well as SEQ ID NOS: 1, 2 and 3), of which the latter is particularly preferred.

According to one specific, but non-limiting aspect of the invention, the amino acid sequence used in the invention is chosen from reductases involved in the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA (i.e. the Second Step above) and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA (i.e. the Fourth Step above). Some preferred, but non-limiting examples of such enzymes are KAR and TER, of which KAR is particularly preferred.

In one more specific embodiment, the amino acid sequence used in the invention may be chosen from the group consisting of ELOVL1, ELOVL2, ELOVL3, ELOVL4, ELOVL5, ELOVL6, KAR and TER, and in particular from ELOVL6, KAR and TER.

Thus, according to a first aspect, the invention relates to a method for identifying a compound that can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases (e.g. from a set or library of test chemicals), said method at least comprising the steps of:

a) contacting an amino acid sequence that is involved in the elongation of fatty acids, and/or a host cell or host organism containing/expressing such an amino acid sequence, with a test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said amino acid sequence; and optionally

b) detecting the signal that may thus be generated, said signal identifying a modulator of said amino acid sequence;

in which the modulator thus identified can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases.

For example, the modulators thus identified can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases, and/or can be used to develop other compounds that can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases, as further described below.

According to a further aspect, the invention relates to a method for generating a signal that is representative for the interaction of a test chemical with an amino acid sequence involved in the elongation of fatty acids, said method at least comprising the steps of:

a) contacting said amino acid sequence, or a host cell or host organism containing/expressing said amino acid sequence involved in the elongation of fatty acids, with said test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said amino acid sequence; and optionally

b) detecting the signal that may thus be generated.

In another aspect, the invention relates to a method for identifying a modulator of an amino acid sequence that is involved in the elongation of fatty acids (e.g. from a set or library of test chemicals), said method at least comprising the steps of:

a) contacting said amino acid sequence, or a host cell or host organism containing/expressing said amino acid sequence, with a test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said amino acid sequence; and optionally

b) detecting the signal that may thus be generated,

wherein a change in said signal (compared to a suitable reference) identifies a modulator of said amino acid sequence.

The "*amino acid sequence involved in elongation of fatty acids*" may be any amino acid sequence that is involved in - e.g. catalyses - the First Step, the Second Step, the Third Step and/or the Fourth Step above in any unicellular or multicellular organism, in particular in yeast or in a multicellular animal, more in particular in a vertebrate animal, even more in particular in a mammal, and specifically in a human

As indicated above, the amino acid sequence involved in elongation of fatty acids is preferably chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

More preferably, the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for fatty acids with a length of the carbon chain of 24 carbon atoms or less; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

Even more preferably, the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for fatty acids with a length of the carbon chain of 20 carbon atoms or less; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

Still more preferably, the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for fatty acids with a length of the carbon chain of 18 carbon atoms or less; and

- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

In particular, the amino acid sequence involved in elongation of fatty acids may be chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for saturated fatty acids with 18 carbon atoms or less in the fatty acid chain, compared to both (1) saturated and unsaturated fatty acids with more than 18 carbon atoms in the fatty acid chain; as well as (2) unsaturated fatty acids of the (n-3) family and/or (n-6) family with 18 carbon atoms or less in the fatty acid chain; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

According to a specific non-limiting aspect, the amino acid sequence involved in elongation of fatty acids may be chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for palmitic acid and/or for stearic acid, compared to both (1) saturated and unsaturated fatty acids with more than 18 carbon atoms in the fatty acid chain; as well as (2) linoleic acid and/or alpha-linolenic acid; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

According to a very specific non-limiting aspect, the amino acid sequence involved in elongation of fatty acids may be chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for palmitic acid,

compared to both (1) stearic acid; as well as (2) linoleic acid and alpha-linolenic acid; and

- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

According to one specific, but non-limiting embodiment, the amino acid sequence corresponding to a condensation enzyme contains at least a histidine-rich motif (H*X*XHH), and preferably also contains an ELO-domain (as determined by SMART™ analysis (see http://smart.embl-heidelberg.de/) and/or preferably by PFAM protein search™ (see for example http://pfam.wustl.edu/).

According to one specific, but non-limiting embodiment, the amino acid sequence corresponding to the reductase involved in the Second Step preferably contains an "*adh_short*" domain (as determined by SMART™ analysis (see http://smart.embl-heidelberg.de/) and/or preferably by PFAM protein search™ (see for example http://pfam.wustl.edu/).

According to one specific, but non-limiting embodiment, the amino acid sequence corresponding to the reductase involved in the Fourth Step preferably contains an "*steriod_short*" domain (as determined by SMART™ analysis (see http://smart.embl-heidelberg.de/) and/or preferably by PFAM protein search™ (see for example http://pfam.wustl.edu/).

According to a more specific but non-limiting embodiment, the amino acid involved in the elongation of fatty acids may be chosen from the group consisting of: ELO1, ELO2, ELO3, *Cig30*, *Ssc1*, *Ssc2*, LCE, ELOVL1, ELOVL2, ELOVL3, ELOVL4, ELOVL5 (HELO1), ELOVL6, KAR and TER, and natural or synthetic analogs thereof, as further defined below.

In particular, the amino acid involved in the elongation of fatty acids may be chosen from the group consisting of: *Cig30*, *Ssc1*, *Ssc2*, LCE, ELOVL1, ELOVL2, ELOVL3, ELOVL4, ELOVL5 (HELO1), ELOVL6, KAR and TER, and natural or synthetic analogs thereof (as further defined below).

More in particular, the amino acid involved in the elongation of fatty acids may be chosen from the group consisting of: LCE, ELOVL6, KAR and TER, and natural or synthetic analogs thereof (as further defined below).

Even more in particular, the amino acid involved in the elongation of fatty acids may be chosen from the group consisting of: ELOVL6, KAR and TER, and natural or synthetic analogs thereof (as further defined below).

Specifically, the amino acid sequence involved in the elongation of fatty acids may be chosen from the group consisting of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33, and natural or synthetic analogs thereof (as further defined below).

More specifically, the amino acid sequence involved in the elongation of fatty acids may be chosen from the group consisting of: SEQ ID NO: 3, 23 and/or 26, and natural or synthetic analogs thereof (as further defined below).

In further aspects, the invention also relates to the use of an amino acid sequence that is involved in the elongation of fatty acids (as defined above), and/or a host cell/host organism that contains and/or expresses such an amino acid sequence sequence, in one of the methods described above. For the purposes herein, the term "*host cell*" also comprises cell fractions and/or preparations derived from such cells, such as cytosolic and in particular microsomal preparations (as mentioned below).

The invention also relates to compounds that can modulate (the biological activity of), and/or that can otherwise interact with, an amino acid that is involved in the elongation of fatty acids. The invention also relates to compositions that contain such compounds, and in particular to pharmaceutical compositions that contain such compounds.

The invention further relates to the use of compounds that can modulate (the biological activity of), and/or that can otherwise interact with, an amino acid sequence that is involved in the elongation of fatty acids, in the preparation of pharmaceutical compositions, and in particular to the use of such compounds in the preparation of a pharmaceutical composition for the prevention and/or treatment of metabolic diseases.

The invention also relates to compounds that can be used in the prevention and/or treatment of metabolic diseases (as further defined below), which compounds have or can be identified and/or developed using an amino acid sequence involved in the elongation

of fatty acids and/or a host cell or host organism that contains and/or expresses such an amino acid sequence. The invention also relates to compositions that contain such compounds, and in particular to pharmaceutical compositions that contain said compounds, and to the use of such compounds in the preparation of a pharmaceutical composition, and in particular to the use of such compounds in the preparation of a pharmaceutical composition for the prevention or treatment of metabolic diseases.

Unless explicitly specified herein, all terms used in the present description have their usual meaning in the art, for which particular reference is made to the definitions given in WO98/06737 and EP 1 085 089.

The amino acid sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33 were identified, and can be derived/isolated from/using human cells; in the manner as further described in the prior art referred to above, or in any other suitable manner known per se.

Generally, the amino acid sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33 may be isolated from the human cells and/or tissues using any technique(s) for protein isolation and/or purification known per se. Alternatively, the amino acid sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33 may be obtained by suitable expression of a suitable nucleotide sequence - such as one of the nucleotide sequences of SEQ ID NOS: 1, 2, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24 or 25, as applicable, or a suitable mutant thereof - in an appropriate host cell or host organism, as further described below.

Also, it is expected that - based upon the disclosure herein - the skilled person will be able to identify, derive and/or isolate natural "analogs" (as mentioned above) of the amino acid sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33. Such mutants could be derived from other human individuals and/or from (individuals of) other species. For example, such natural analogs may be obtained from yeast or from species of multicellular animals such as *Drosophila melanogaster* or *Caenorhabditis elegans*, and in particular from other species of vertebrate animals, more in particular of mammals, such as rat, mouse, pig, sheep, rabbit, cow, horse or dog.

Such natural analogs may again be obtained by isolating them from their natural source using any technique(s) for protein isolation and/or purification known per se, or

alternatively by suitable expression of a suitable nucleotide sequence - such as a natural "mutant" as described herein - in an appropriate host cell or host organism, as further described below.

It is also expected that - based upon the disclosure herein - the skilled person will be able to provide and/or derive synthetic "analogs" (as mentioned above) of one or more of the amino sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33.

Generally, such synthetic analogs may be obtained by suitable expression of a suitable nucleotide sequence used in the invention - such as a synthetic mutant as described above - in an appropriate host cell or host organism, as further described below.

Preferably, any analogs as described herein will have one or more, and preferably all, of the structural characteristics/conserved features referred to above for the sequences of SEQ ID NO: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33, such as - in the case of a condensation enzyme involved in the First Step above - the HXXHH motif and/or the ELO domain, and – in the case of a reductase involved in the Second Step above – an adh_short domain.

It is also possible in the invention to use a part or fragment of one or more of the amino acid sequences of SEQ ID NOS 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33, or a part or fragment of a (natural or synthetic) analog thereof. This may for instance be a N- and/or C- truncated amino acid sequence. Also, two or more parts or fragments of one or more amino acid sequences involved in the elongation of fatty acids may be suitably combined to provide a (further) amino acid sequence for use in the invention.

Preferably, any such parts or fragments will be such that they comprise at least one continuous stretch of at least 5 amino acids, preferably at least 10 amino acids, more preferably at least 20 amino acids, even more preferably more than 30 amino acids, of one or more of the amino acid sequences of SEQ ID NO: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33.

In particular, any parts or fragments as described herein are such that they (at least) comprise the active/catalytic site of the corresponding amino acid sequence involved in the elongation of fatty acids and/or a binding domain of the corresponding amino acid sequence involved in the elongation of fatty acids, such as - in the case of a condensation enzyme involved in the First Step above - the HXXHH motif and/or (part

of) the ELO domain, and - in the case of a reductase involved in the Second Step above - an adh_short domain. As will be clear to the skilled person, such parts or fragments may find particular use in assay- and screening techniques (as generally described below) and/or (when said part or fragment is provided in crystalline form) in X-ray crystallography.

Generally, such parts or fragments of the amino acid sequences involved in the elongation of fatty acids may be obtained by suitable expression of a suitable nucleotide sequence used in the invention - such as a suitable part or fragment as described herein for the nucleotide sequences used in the invention - in an appropriate host cell or host organism, as further described below.

In addition and/or as an alternative to the methodology above, amino acid sequences used in the invention may also be provided by (chemically and/or enzymatically) modifying the side chain(s) of one or more amino acid residues of an amino acid sequence of SEQ ID NO: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33 or a part, fragment, (natural and/or synthetic) mutant, variant, allele, analogs, orthologs thereof, for example by one or more of the side chain modifications as described in WO 01/02560 and/or by incorporating (e.g. by insertion and/or substitution) one or more unnatural amino acid residues, again as described in WO 01/02560.

Preferably, any analogs, parts and/or fragments as used herein will be such that they have a degree of "sequence identity", at the amino acid level, with one or more of the amino acid sequences of SEQ ID NOS 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33 of at least 50%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95 % or more.

For this purpose, the percentage of "sequence identity" between a given amino acid sequence and one of the amino acid sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33 may be calculated by dividing [*the number of amino acid residues in the given amino acid sequence that are identical to the amino acid residue at the corresponding position in the amino acid sequence of the relevant SEQ ID NO]* by [*the total number of amino acid residues in the given amino acid sequence*] and multiplying by [*100%*], in which each deletion, insertion, substitution or addition of an amino acid

residue - compared to the sequence of the relevant SEQ ID NO - is considered as a difference at a single amino acid (position).

Alternatively, the degree of sequence identity may be calculated using a known computer algorithm, such as those mentioned above.

Also, such sequence identity at the amino acid level may take into account so-called "conservative amino acid substitutions", which are well known in the art, for example from GB-A-2 357 768, WO 98/49185, WO 00/46383 and WO 01/09300; and (preferred) types and/or combinations of such substitutions may be selected on the basis of the pertinent teachings from the references mentioned in WO 98/49185.

Also, preferably, any analogs, parts and/or fragments as used herein will have a biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33, i.e. to a degree of at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90%, as measured by a suitable assay method, for example those mentioned in the prior art cited herein for each of these sequences.

Preferably, an amino acid sequence used in the invention will (also) have a length (expressed as total number of amino acid residues), which is at least 50%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95 % or more of the length of one or more of the amino acid sequence of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33.

It is also within the scope of the invention to use a fusion of an amino acid sequence involved in the elongation of fatty acids (as described above) with one or more further amino acid sequences, for example to provide a protein fusion. Generally, such fusions may be obtained by suitable expression of a suitable nucleotide sequence used in the invention - such as a suitable fusion of a nucleotide sequence used in the invention with one or more further coding sequences - in an appropriate host cell or host organism, as further described below.

One particular embodiment, such fusions may comprise an amino acid sequence involved in the elongation of fatty acids fused with a reporter protein such as GFP, luciferase or another fluorescent protein moiety. As will be clear to the skilled person, such fusions may find particular use in expression analysis and similar methodologies.

In another embodiment, the fusion partner may be an amino acid sequence or residue that may be used in purification of the expressed amino acid sequence, for example using affinity techniques directed against said sequence or residue. Thereafter, said sequence or residue may be removed (e.g. by chemical or enzymatical cleavage) to provide the nucleotide sequence used in the invention (for this purpose, the sequence or residue may optionally be linked to the amino acid sequence involved in the elongation of fatty acids via a cleavable linker sequence). Some preferred, but non-limiting examples of such residues are multiple histidine residues and glutatione residues,

In one preferred, but non-limiting aspect, any such fusion will have a biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33, i.e. to a degree of at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90%, as measured by a suitable assay method, for example those mentioned in the prior art cited herein for each of these sequences.

As mentioned, the amino acid sequences involved in the elongation of fatty acids, or natural or synthetic analogs thereof, may be obtained by suitable expression of a nucleotide sequence encoding such an amino acid sequence in a suitable host organism, as further described below. According to one embodiment, said nucleotide sequence may be in the form of a genetic construct, also as further described below.

Some preferred, but non-limiting examples of nucleotide sequences that can be used to express the amino acid sequences involved in the elongation of fatty acids include, but are not limited to, the nucleotide sequences of SEQ ID NOS: 1, 2, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24 or 25, as applicable, or a suitable part, mutant, variant, allele, analog, orthologs, fusion or splice variant thereof (collectively referred to herein as "*mutants*"), depending on the desired amino acid sequence.

The nucleotide sequences of SEQ ID NOS: 1, 2, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24 or 25 were identified, and can be derived/isolated from/using human cells and/or tissues; in any suitable manner known per se, including but not limited to PCR starting from human genomic DNA or a library of human cDNA, using primers designed on the basis of the relevant sequence.

Also, it is expected that - based upon the disclosure herein - the skilled person will be able to identify, derive and/or isolate natural "mutants" (as mentioned above) of the above nucleotide sequences. For example, such mutants could be derived from other species (in which case these mutants will also be referred to herein as "*orthologs*"). Some examples of species from which such orthologs could be derived include, but are not limited to species of

- unicellular and/or micro-organisms such as bacteria, and yeast,
- invertebrate multicellular organisms as such as insects and nematodes (for example, agronomically harmful insect or nematode species);
- vertebrate multicellular organisms as such as fish, birds, reptiles, amphibians and mammals;

Preferably, a natural ortholog is derived from a mammal such as a mouse, rat, rabbit or dog and the further species mentioned above..

Such natural mutants may be obtained in a manner essentially analogous to the methods described in the prior art referred to above, or alternatively by:

- construction of a DNA library from the species of interest in an appropriate expression vector system, followed by direct expression of the mutant sequence;
- construction of a DNA library from the species of interest in an appropriate expression vector system, followed by screening of said library with a probe used in the invention (as described below) and/or with a(nother) nucleotide sequence used in the invention;
- isolation of mRNA that encodes the mutant sequence from the species of interest, followed by cDNA synthesis using reverse transcriptase;

and/or by any other suitable method(s) or technique(s) known per se, for which reference is for instance made to the standard handbooks, such as Sambrook et al, "Molecular Cloning: A Laboratory Manual" (2nd.ed.), Vols. 1-3, Cold Spring Harbor Laboratory Press (1989) and F. Ausubel et al, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987).

It is also expected that - based upon the disclosure herein - the skilled person will be able to provide and/or derive synthetic mutants (as defined hereinabove) of the above nucleotide sequences.

Techniques for generating such synthetic sequences will be clear to the skilled person and may for instance include, but are not limited to, automated DNA synthesis; site-directed mutagenesis; combining two or more parts of one or more naturally occurring sequences, introduction of mutations that lead to the expression of a truncated expression product; introduction of one or more restriction sites (e.g. to create casettes and/or regions that may easily be digested and/or ligated using suitable restriction enzymes), and/or the introduction of mutations by means of a PCR reaction using one or more "mismatched" primers, using for example a sequence of a naturally occurring GPCR as a template. These and other techniques will be clear to the skilled person, and reference is again made to the standard handbooks, such as Sambrook et al. and Ausubel et al., mentioned above.

Preferably, any mutants used will encode amino acid sequences having one or more, and preferably all, of the structural characteristics/conserved features referred to for the elongation enzymes above (such as - in the case of a condensation enzyme involved in the First Step above - an HXXHH motif and/or an ELO domain in the case of an enzyme involved in the condensation reaction of the First Step above, and - in the case of a reductase involved in the Second Step above - an adh_short domain) and in particular the active site(s) and/or catalytic domain(s).

According to one embodiment, the "mutants" used are (also) such that they are capable of hybridizing with one or more of the nucleotide sequences of SEQ ID NO's 1, 2, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24 or 25, i.e. under conditions of "moderate stringency", and preferably under conditions of "high stringency". Such conditions will be clear to the skilled person, for example from the standard handbooks, such as Sambrook et al. and Ausubel et al., mentioned above, as well as in EP 0 967 284, EP 1 085 089 or WO 00/55318; particular reference is made to the "stringent" hybridisation conditions described in WO 00/78972 and WO 98/49185, and/or the hybridization conditions described in GB 2 357 768-A.

Also, any "mutants" used will preferably have a degree of "sequence identity", at the nucleotide level, with one or more of the nucleotide sequences of SEQ ID NOS: 1, 2, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24 or 25 of at least 50%, preferably at least

60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90%, and up to 95% or more.

For this purpose, the percentage of "sequence identity" between a given nucleotide sequence and one of the nucleotide sequences used in the invention may be calculated by dividing [*the number of nucleotides in the given nucleotide sequence that are identical to the nucleotide at the corresponding position in the nucleotide sequence of the relevant SEQ ID NO*] by [*the total number of nucleotides in the given nucleotide sequence*] and multiplying by [*100%*], in which each deletion, insertion, substitution or addition of a nucleotide - compared to the sequence of the relevant SEQ ID NO - is considered as a difference at a single nucleotide (position).

Alternatively, the degree of sequence identity may be calculated using a known computer algorithm for sequence alignment such as NCBI Blast v2.0, using standard settings.

Some other techniques, computer algorithms and settings for determining the degree of sequence identity are for example described in EP 0 967 284, EP 1 085 089, WO 00/55318, WO 00/78972, WO 98/49185 and GB 2 357 768-A.

Generally, the nucleotide sequences used in the invention, when in the form of a nucleic acid, may be DNA or RNA, and may be single stranded or double stranded. For example, the nucleotide sequences used in the invention may be genomic DNA, cDNA or synthetic DNA (such as DNA with a codon usage that has been specifically adapted for expression in the intended host cell or host organism). Thus, the nucleotide sequences used in the invention may contain intron sequences, and also generally comprises different splice variants.

It is also within the scope used in the invention to use a fusion of a nucleotide sequence used in the invention (as described above) with one or more further nucleotide sequence(s), including but not limited to one or more coding sequences, non-coding sequences and/or regulatory sequences. Preferably, in such fusions, the one or more further nucleotide sequences are operably connected (as described below) to the nucleotide sequence used in the invention (for example so that, when the further nucleotide sequence is a coding sequence, the nucleotide fusion encodes a protein fusion as described below).

Another embodiment used in the invention relates to a nucleic acid probe that is capable of hybridizing with a nucleotide sequence of SED ID NO: 1, 2, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24 or 25 under conditions of moderate stringency, preferably under conditions of high stringency, and in particular under stringent conditions (all as described above). Such nucleotide probes may for instance be used for detecting and/or isolating a(nother) nucleotide sequence used in the invention and/or as a primer for amplifying a nucleotide sequence used in the invention; all using techniques known per se, for which reference is again made to the general handbooks such as Sambrook et al. and Ausubel et al. mentioned above.

Generally, such probes can be designed by the skilled person starting from a nucleotide sequence and/or amino acid sequence used in the invention - and in particular one or more of the sequences of SEQ ID NOS 1, 2, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24 or 25.

The nucleic acids used in the invention may also be in the form of a genetic construct, which may for example comprise:

a) the nucleotide sequence as described above; operably connected to:

b) one or more regulatory elements, such as a promoter and optionally a suitable terminator;

and optionally also:

c) one or more further elements of genetic constructs known per se;

in which the terms "*regulatory element*", "*promoter*", "*terminator*", "*further elements*" and "*operably connected*" have the meanings indicated hereinbelow.

Such genetic constructs may be DNA or RNA, and are preferably double-stranded DNA. The constructs may also be in a form suitable for transformation of the intended host cell or host organism, in a form suitable for integration into the genomic DNA of the intended host cell or in a form suitable independent replication, maintenance and/or inheritance in the intended host organism. For instance, the genetic construct may be in the form of a vector, such as for example a plasmid, cosmid, YAC, a viral vector or transposon. In particular, the vector may be an expression vector, i.e. a vector that can provide for expression *in vitro* and/or *in vivo* (e.g. in a suitable host cell and/or host organism as described below).

As the one or more "further elements" referred to above, the genetic construct(s) used in the invention may generally contain one or more suitable regulatory elements (such as a suitable promoter(s), enhancer(s), terminator(s), etc.), 3'- or 5'-UTR sequences, leader sequences, selection markers, expression markers/reporter genes, and/or elements that may facilitate or increase (the efficiency of) transformation or integration. These and other suitable elements for such genetic constructs will be clear to the skilled person, and may for instance depend upon the type of construct used, the intended host cell or host organism; the manner in which the nucleotide sequences used in the invention of interest are to be expressed (e.g. via constitutive, transient or inducible expression); and/or the transformation technique to be used.

Preferably, in the genetic constructs used in the invention, the one or more further elements are "*operably linked*" to the nucleotide sequence(s) used in the invention and/or to each other, by which is generally meant that they are in a functional relationship with each other. For instance, a promoter is considered "*operably linked*" to a coding sequence if said promoter is able to initiate or otherwise control/regulate the transcription and/or the expression of a coding sequence (in which said coding sequence should be understood as being "*under the control of*" said promotor)

Generally, when two nucleotide sequences are operably linked, they will be in the same orientation and usually also in the same reading frame. They will usually also be essentially contiguous, although this may also not be required.

Preferably, the optional further elements of the genetic construct(s) used in the invention are such that they are capable of providing their intended biological function in the intended host cell or host organism.

For instance, a promoter, enhancer or terminator should be "*operable*" in the intended host cell or host organism, by which is meant that (for example) said promoter should be capable of initiating or otherwise controlling/regulating the transcription and/or the expression of a nucleotide sequence - e.g. a coding sequence - to which it is operably linked (as defined above).

Such a promoter may be a constitutive promoter or an inducible promoter, and may also be such that it (only) provides for expression in a specific stage of development

of the host cell or host organism, and/or such that it (only) provides for expression in a specific cell, tissue, organ or part of a multicellular host organism.

Some particularly preferred promoters include, but are not limited to those present in the expression vectors referred to below.

A selection marker should be such that it allows - i.e. under appropriate selection conditions - host cells and/or host organisms that have been (successfully) transformed with the nucleotide sequence used in the invention to be distinguished from host cells/organisms that have not been (successfully) transformed. Some preferred, but non-limiting examples of such markers are genes that provide resistance against antibiotics (such as kanamycine or ampicilline), genes that provide for temperature resistance, or genes that allow the host cell or host organism to be maintained in the absence of certain factors, compounds and/or (food) components in the medium that are essential for survival of the non-transformed cells or organisms.

A leader sequence should be such that - in the intended host cell or host organism - it allows for the desired post-translational modifications and/or such that it directs the transcribed mRNA to a desired part or organelle of a cell. A leader sequence may also allow for secretion of the expression product from said cell. As such, the leader sequence may be any pro-, pre-, or prepro-sequence operable in the host cell or host organism.

An expression marker or reporter gene should be such that - in the host cell or host organism - it allows for detection of the expression of (a gene or nucleotide sequence present on) the genetic construct. An expression marker may optionally also allow for the localisation of the expressed product, e.g. in a specific part or organelle of a cell and/or in (a) specific cell(s), tissue(s), organ(s) or part(s) of a multicellular organism. Such reporter genes may also be expressed as a protein fusion with the amino acid sequence involved in the elongation of fatty acids. Some preferred, but non-limiting examples include fluorescent proteins such as GFP.

For some (further) non-limiting examples of the promoters, selection markers, leader sequences, expression markers and further elements that may be present/used in the genetic constructs used in the invention - such as terminators, transcriptional and/or translational enhancers and/or integration factors - reference is made to the general handbooks such as Sambrook et al. and Ausubel et al. mentioned above, to W.B. Wood et

al., "*The nematode Caenorhabditis elegans*", Cold Spring Harbor Laboratory Press (1988) and D.L. Riddle et al., "*C. ELEGANS II*", Cold Spring Harbor Laboratory Press (1997), as well as to the examples that are given in WO 95/07463, WO 96/23810, WO 95/07463, WO 95/21191, WO 97/11094, WO 97/42320, WO 98/06737, WO 98/21355, US-A-6,207,410, US-A- 5,693,492 and EP 1 085 089. Other examples will be clear to the skilled person.

The genetic constructs used in the invention may generally be provided by suitably linking the nucleotide sequence(s) used in the invention to the one or more further elements described above, for example using the techniques described in the general handbooks such as Sambrook et al. and Ausubel et al., mentioned above.

Often, the genetic constructs used in the invention will be obtained by inserting a nucleotide sequence used in the invention in a suitable (expression) vector known per se. Some preferred, but non-limiting examples of suitable expression vectors include:

- vectors for expression in mammalian cells: pMAMneo (Clontech), pcDNA3 (Invitrogen), pMC1neo (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593), pBPV-1 (8-2) (ATCC 37110), pdBPV-MMTneo (342-12) (ATCC 37224), pRSVgpt (ATCC37199), pRSVneo (ATCC37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460) and 1ZD35 (ATCC 37565);
- vectors for expression in bacterials cells: pET vectors (Novagen) and pQE vectors (Qiagen);
- vectors for expression in yeast or other fungal cells: pYES2 (Invitrogen) and Pichia expression vectors (Invitrogen);
- vectors for expression in insect cells: pBlueBacII (Invitrogen).

The nucleotide sequences and/or genetic constructs used in the invention may be used to transform a host cell or host organism.

The host cell may be any suitable (fungal, prokaryotic or eukaryotic) cell or cell line, for example:

- a bacterial strain, including but not limited to strains of *E.coli*, *Bacillus*. *Streptomyces* and *Pseudomonas*;
- a fungal cell, including but not limited to cells from species of *Aspergillus* and *Trichoderma*;

- a yeast cell, including but not limited to cells from species of *Kluyveromyces* or *Saccharomyces;*
- an amphibian cell or cell line, such as *Xenopus* oocytes.

In one specific embodiment, which may particularly useful when the nucleotide sequences used in the invention are (to be) used in the discovery and development of insecticidal compounds, the host cell may be an insect-derived cell or cell line, such as:

- cells/cell lines derived from *lepidoptera*, including but not limited to *Spodoptera* SF9 and Sf21 cells,
- cells/cell lines derived from *Drosophila*, such as Schneider and Kc cells; and/or
- cells/cell lines derived from a pest species of interest (as mentioned below), such as from *Heliothis virescens*.

In one preferred embodiment, the host cell is a mammalian cell or cell line, for example derived from the mammals referred to above.

In an even more preferred aspect, the host cell is a cell or cell line derived from a human, from the mammals including but not limited to CHO- and BHK-cells and human cells or cell lines such as HeLa and COS.

In one specific, but non-limiting embodiment, the cell or cell line may be human cell or cell line which is related to metabolic processes or metabolic disease and/or used as a cellular model for metabolic disease, including but not limited to liver cells or cell lines, adipocytes or muscle cells or cell lines such as HEPG2 cells, 3T3L1 adipocytes, CTC12 cells and L6 myotubes.

The host organism may be any suitable multicellular (vertebrate or invertebrate) organism, including but not limited to:

- a nematode, including but not limited to nematodes from the genus *Caenorhabditis*, such as *C.elegans*,
- an insect, including but not limited to species of *Drosophila* and/or a specific pest species of interest (such as those mentioned above);
- other well known model organisms, such as zebrafish;
- a mammal such as a rat or mouse;

Other suitable host cells or host organisms will be clear to the skilled person, for example from the handbooks and patent applications mentioned above.

It should be noted that when a nucleotide sequence used in the invention is expressed in a multicellular organism, it may be expressed throughout the entire organism, or only in one or more specific cells, tissues, organs and/or parts thereof, for example by expression under the control of a promoter that is specific for said cell(s), tissue(s), organ(s) or part(s).

The nucleotide sequence may also be expressed during only a specific stage of development or life cycle of the host cell or host organism, again for example by expression under the control of a promoter that is specific for said stage of development or life cycle. Also, as already mentioned above, said expression may be constitutive, transient and/or inducible.

According to one specific embodiment, the expression of a nucleotide sequence used in the invention in a host cell or host organism may be party or totally reduced (i.e. knocked out), compared to the original (e.g. native) host cell or host organism. This may for instance be achieved in a transient manner using antisense and/or RNA-interference techniques well known in the art, or in a constitutive manner using random, site specific and/or chemical mutagenesis of the nucleotide sequence used in the invention, or any other suitable techniques for generating "knock-down" or "knock-out" animals.

Suitable transformation techniques will be clear to the skilled person and may depend on the intended host cell/host organism and the genetic construct to be used. Some preferred, but non-limiting examples of suitable techniques include ballistic transformation, (micro-)injection, transfection (e.g. using suitable transposons), electroporation and lipofection. For these and other suitable techniques, reference is again made to the handbooks and patent applications mentioned above.

After transformation, a step for detecting and selecting those host cells or host organisms that have been succesfully transformed with the nucleotide sequence/genetic construct used in the invention may be performed. This may for instance be a selection step based on a selectable marker present in the genetic construct used in the invention or a step involving the detection of the amino acid sequence involved in the elongation of fatty acids, e.g. using specific antibodies.

The transformed host cell (which may be in the form or a stable cell line) or host organisms (which may be in the form of a stable mutant line or strain) form further aspects of the present invention.

Preferably, these host cells or host organisms are such that they express, or are (at least) capable of expressing (e.g. under suitable conditions), an amino acid sequence involved in the elongation of fatty acids (and in case of a host organism: in at least one cell, part, tissue or organ thereof). The invention also includes further generations, progeny and/or offspring of the host cell or host organism used in the invention, that may for instance be obtained by cell division or by sexual or asexual reproduction.

To produce/obtain expression of the amino acid sequences used in the invention, the transformed host cell or transformed host organism may generally be kept, maintained and/or cultured under conditions such that the (desired) amino acid sequence involved in the elongation of fatty acids is expressed/produced. Suitable conditions will be clear to the skilled person and will usually depend upon the host cell/host organism used, as well as on the regulatory elements that control the expression of the (relevant) nucleotide sequence used in the invention. Again, reference is made to the handbooks and patent applications mentioned above in the paragraphs on the genetic constructs used in the invention.

Generally, suitable conditions may include the use of a suitable medium, the presence of a suitable source of food and/or suitable nutrients, the use of a suitable temperature, and optionally the presence of a suitable inducing factor or compound (e.g. when the nucleotide sequences used in the invention are under the control of an inducible promoter); all of which may be selected by the skilled person. Again, under such conditions, the amino acid sequences involved in the elongation of fatty acids may be expressed in a constitutive manner, in a transient manner, or only when suitably induced.

It will also be clear to the skilled person that the amino acid sequence involved in the elongation of fatty acids may (first) be generated in an immature form (as mentioned above), which may then be subjected to post-translational modification, depending on the host cell/host organism used. Also, the amino acid sequence involved in the elongation of fatty acids may be glycosylated, again depending on the host cell/host organism used.

The amino acid sequence involved in the elongation of fatty acids may then be isolated from the host cell/host organism and/or from the medium in which said host cell or host organism was cultivated, using protein isolation and/or purification techniques known per se, such as (preparative) chromatography and/or electrophoresis techniques, differential precipitation techniques, affinity techniques (e.g. using a specific, cleavable amino acid sequence fused with the amino acid sequence involved in the elongation of fatty acids) and/or preparative immunological techniques (i.e. using antibodies against the amino acid sequence to be isolated).

In one embodiment, the amino acid sequence thus obtained may also be used to generate antibodies specifically against said sequence or an antigenic part or epitope thereof.

Such antibodies, which form a further aspect of the invention, may be generated in a manner known per se, for example as described in GB-A-2 357 768, US-A-5,693,492, WO 95/32734, WO 96/23882, WO 98/02456, WO 98/41633 and/or WO 98/49306, and/or as described in the prior art referred to above. Often, but not exclusively, such methods will involve immunizing a immunocompetent host with the pertinent amino acid sequence involved in the elongation of fatty acids or an immungenic part thereof (such as a specific epitope), in amount(s) and according to a regimen such that antibodies against said amino acid sequence are raised, and than harvesting the antibodies thus generated, e.g. from blood or serum derived from said host.

For instance, polyclonal antibodies can be obtained by immunizing a suitable host such as a goat, rabbit, sheep, rat, pig or mouse with (an epitope of) an amino acid sequence involved in the elongation of fatty acids, optionally with the use of an immunogenic carrier (such as bovine serum albumin or keyhole limpet hemocyanin) and/or an adjuvant such as Freund's, saponin, ISCOM's, aluminium hydroxide or a similar mineral gel, or keyhole limpet hemocyanin or a similar surface active substance. After a suitable immune response has been raised (usually within 1-7 days), the antibodies can be isolated from blood or serum taken from the immunized animal in a manner known per se, which optionally may involve a step of screening for an antibody with desired properties (i.e. specificity) using known immunoassay techniques, for which reference is again made to for instance WO 96/23882.

Monoclonal antibodies may for example be produced using continuous cell lines in culture, including hybridoma-based and similar techniques, again essentially as described in the above cited references. Accordingly, cells and cell lines that produce monoclonal antibodies against an amino acid sequence involved in the elongation of fatty acids form a further aspect of the invention, as do methods for producing antibodies against amino acid sequences involved in the elongation of fatty acids, which methods may generally involve cultivating such a cell and isolating the antibodies from the culture (medium), again using techniques known per se.

Also, Fab-fragments against the amino acid sequences involved in the elongation of fatty acids (such as $F(ab)_2$, Fab' and Fab fragments) may be obtained by digestion of an antibody with pepsin or another protease, reducing disulfide-linkages and treatment with papain and a reducing agent, respectively. Fab-expression libraries may for instance be obtained by the method of Huse et al., 1989, Science 245:1275-1281.

In another embodiment, the nucleotide sequences used in the invention, the amino acid sequences used in the invention, and/or a host cell or host organism that expresses such an amino acid sequence, may also be used in an assay or assay method generally (including but not limited to diagnostic assays and/or assays to determing the presense and/or absence of specific mutations and/or genetic markers, for example to determine susceptibility for a condition or disease associated with such a mutation or marker), and in particular in an assay to identify and/or (further) develop compounds and/or other factors that can modulate the (biological) activity of, and/or that can otherwise interact with, the amino acid sequences involved in the elongation of fatty acids, and such uses form further aspects of the invention. As will be clear to the skilled person, in this context, the amino acid sequence involved in the elongation of fatty acids will serve as a target for interaction with such a compound or factor

In this context, the terms "*modulate*", "*modulation,* "*modulator*" and "*target*" will have their usual meaning in the art, for which reference is *inter alia* made to the definitions given in WO 98/06737. Generally, a modulator is a compound or factor that can enhance, inhibit/reduce or otherwise alter, influence or affect (collectively referred to as "*modulation*") a functional property of a biological activity or process (for example,

the biological activity of an amino acid sequence involved in the elongation of fatty acids).

In this context, the amino acid sequence involved in the elongation of fatty acids may serve as a target for modulation *in vitro* (e.g. as part of an assay or screen) and/or for modulation *in vivo* (e.g. for modulation by a compound or factor that is known to modulate the target, which compound or factor may for example be used as an active compound for agrochemical, veterinary and/or pharmaceutical use).

For example, the amino acid sequences, host cells and/or host organisms used in the invention may be used as part of an assay or screen that may be used to identify and/or develop modulators of the amino acid sequence involved in the elongation of fatty acids, such as a primary screen (e.g. a screen used to identify modulators of the target from a set or library of test chemicals with unknown activity with respect to the target) and/or a secondary assay (e.g. an assay used for validating hits from a primary screen and/or used in optimizing hit molecules, e.g. as part of hits-to-leads chemistry).

For instance, such an assay or screen may be configured as an *in vitro* assay or screen, which will generally involve binding of the compound or factor to be tested as a potential modulator for the target (hereinbelow also referred to as "test chemical") to the target, upon which a signal generated by said binding is measured. Suitable techniques for such *in vitro* screening will be clear to the skilled person, and are for example described in Eldefrawi et al., (1987). FASEB J., Vol.1, pages 262-271 and Rauh et al., (1990), Trends in Pharmacol. Sci., vol.11, pages 325-329. For example, such an assay or screen may be configured as a binding assay or screen, in which the test chemical is used to displace a detectable ligand from the target (e.g. a radioactive or fluorescent ligand), upon which the amount of ligand displaced from the target by the modulator is determined. Other suitable assays for the amino acid sequences involved in the elongation of fatty acids will be clear to the skilled person, and may for example be found in the prior art cited herein; such assays may optionally also be adapted to and/or configured for screening in an automated, medium-to-high throughput fashion.

For example, reductases involved the elongation of fatty acids such as KAR and TER may be screened for modulators (such as inhibitors) using one of the following techniques:

1) Radioactive detection method based on the use of a radioactive labeled substrate, for example ^{14}C labeled Malonyl-CoA (see also Moon and Horton, *supra*). To achieve the separation between the reactants and products that facilitates quantification (in particular when the assay is performed in a high throughput mode, compared to the assay of Moon and Horton), the fatty acid chain (e.g. palmitoyl-CoA,) could be bound to streptavidin (directly or via biotin) instead of BSA (as used by Moon and Horton), which in turn would be coated on the bottom of the microtitre well plate. Upon the completion and stopping of the enzymatic reaction, the plates would be subjected to a washing protocol prior to the scintillation measurement. The quantity if the incorporated ^{14}C would be proportional to the quantity of the ^{14}C malonyl-CoA condensed to the palmytoyl-CoA.

   Alternatively, an SPA™ assay can be envisaged in which the palmitoyl-CoA is attached to an SPA bead. Upon the reaction with ^{14}C-labeled malonyl-CoA, the ^{14}C would be incorporated to the fatty acid chain fixed to the SPA bead, generating scintillation from the bead due to the proximity of the radioactive label. (Amersham patented technology). The advantage of this assay configuration would be that no washing is required.

2) Luminescent assay based on the detection of FMN cofactor (see for example Warne, et al, 2002, Poster #2448, 8th Annual Conference and Exhibition of the Society for Biomolecular Screening, September 22-26th 2002, The Hague, the Netherlands), which is suitable for the detection of the activity of all the three enzymes involved in the fatty acid elongation. This may be performed either by using the isolated enzyme, or a suitable microsomal preparation containing tow or more, or essentially all, enzymes of the elongation pathway.

3) Colorimetric detection of NADH consumption. In the assay described by Moon and Horton, *supra*, NADH is used as a source of hydrogen in the reduction of the 3-ketoacyl-CoA. A colorimetric method has been described recently to detect directly the consumption of NADH and formation of NAD (See Mayer and Arnold, 2002, J. Biomol. Screening, Vol.7 (2), pp135). In this way, the reduction reaction catalyzed by KAR can be monitored by measuring the consumption of NADH.

4) Fluorescence Resonance Electron Transfer (FRET) can be detected when two suitable fluorescent molecules are brought in proximity of each other. As the enzymatic reaction performed by KAR is a reduction, it should be in principle possible to detect product formation by FRET between the individually labeled reactants, provided the the fluorophore are chosen such (especially respective to the size of malonyl-CoA) that steric hindrances can avoided;

and/or by one of the assays described in the prior art for KAR and TER, such as the prior art mentioned herein, which assay may optionally be adapted and/or configured for automation and/or medium to high through-put

The condensation enzymes involved in the First Step above could also be screened using a luminescent assay based on the detection of FMN cofactor, as described under 2) above; and/or by one of the assays described in the prior art for said condensation enzymes, such as the prior art mentioned herein, which assay may optionally be adapted and/or configured for automation and/or medium to high through-put

Assays or screens for identifying compounds that can interact with the amino acid sequences involved in the elongation of fatty acids may also be configured as a cell-based assay or screen, in which a host cell used in the invention is contacted with/exposed to a test chemical, upon which at least one biological response by the host cell is measured. For example, such cell-based assays may be performed by (over)expression of cDNA encoding the elongase(s) to be screened in *S. cerevisiae* (as performed for the human elongases HELO1 (= ELOVL5) by Leonard et al., *supra*) or in a mutant of *S. cerevisiae* that lacks one or more of the native elongations enzymes ELO1, ELO2 and ELO3 (see Toke et al. and Oh et al., both *supra*). The influence of the compound on the elongase(s) may then be determined by exposing the cells to a suitable substrate and the compound(s) to be tested, analogous to the methods described in the references just cited.

Such cell-based methods may also allow, with advantage, to screen the influence of the compound(s) to be tested not only on the elongase(s) to be screened, but also on the whole elongation pathway or cell, and may as such form a suitable secondary assay for confirmation of hits generated using an in vitro HTS screen as described above.

Alternatively, for identifying compounds that influence the entire elongation pathway (i.e. two or more enzymes therein), the luminescent assay based on the detection of FMN cofactor described under 2) above could be used, using a suitable microsomal preparation containing all enzymes from the elongation pathway.

Suitable cells or cell lines for such cell based assays include those mentioned above. In one preferred, but non-limitig embodiment, the cell or cell line may be a mammalian, and in particular human, cell or cell line which is related to metabolic processes or metabolic disease and/or used as a cellular model for metabolic disease, including but not limited to liver cells or cell lines, adipocytes or muscle cells or cell lines such as HEPG2 cells, 3T3L1 adipocytes, CTC12 cells and L6 myotubes.

Also, such an assay or screen may also be configured as an whole animal screen, in which a host organism used in the invention is contacted with/exposed to a test chemical, upon which at least one biological response (such as a phenotypical, behavioural and/or fysiological change, including but not limited to paralysis or death) by the host organism is measured. Such screens may be carried out in any model organism known per se, including but not limited to yeast, Drosophila, zebrafish or *C. elegans*.

Thus, generally, the assays and screens described above will comprise at least one step in which the test chemical is contacted with the target (and/or with a host cell or host organism that expresses the target), and in particular in such a way that a signal is generated that is representative for the modulation of the target by the test chemical. In a further step, said signal may then be detected.

The compounds that may be tested using the methods of the invention are generally described below.

The assays and screens used in the invention may be carried out at medium throughput to high throughput, for example in an automated fashion using suitable robotics. In particular, in this embodiment, the method used in the invention may be carried out by contacting the target with the test compound in a well of a multi-well plate, such as a standard 24, 96, 384, 1536 or 3456 well plate.

Usually, in a screen or assay used in the invention, for each measurement, the target or host cell/host organism will be contacted with only a single test compound. However, it is also within the scope of the invention to contact the target with two or

more test compounds - either simultaneously or sequentially - for example to determine whether said combination provides a synergistic effect.

According to one specific, but non-limiting embodiment of the invention, when the compound is intended to interact with a condensation enzyme involved in the First Step above, the compound used is also preferably such that it is selective for one condensation enzyme (i.e. with a specific substrate specificity, as described above) compared to another condensation enzyme (i.e. with a different substrate specificity).

In particular, the compound used is preferably such that it is selective for the condensation enzymes which are indicated hereinabove as preferred, compared to condensation enzymes that indicated hereinabove as being less preferred. For example, the compound used is preferably selective for a condensation enzyme that has specificity (as defined above) for fatty acids with a length of the carbon chain of 20 carbon atoms or less, compared to a condensation enzyme that has specificity for fatty acids with a length of the carbon chain of more than 20 carbon atoms.

In particular, the compound used is preferably such that is has specificity for LCE and/or ELOVL6, compared to other condensation enzymes.

The selectivity of a compound for one enzyme/isoform compared to another may be determined in a manner known per se, which generally involves comparing the influence the compound has on the activity of one enzyme/isoform compared to another, using a suitable assay (usually at the same concentration(s) of compound, enzyme and substrate). For example, when the compound is an inhibitor, the selectivity may be established by comparing the IC50 or ED50 values of the compound with respect of each enzyme, as is well known in the art, with a lower ED50 or IC50 value with respect to a first enzyme compared to a second enzyme being an indication that the compound is more selective for the first enzyme compared to the second.

In particular, in the invention, a compound is considered selective for a first enzyme compared to a second when the ED50 or IC50 value of the compound with respect to the first enzyme is less than 90%, preferably less than 50%, more preferably less than 25%, even more preferably less than 10%, an in particular less than 5%, such as less than 1% or even less than 0.1%, of the ED50 or IC50 value for the second enzyme.

Once a test chemical has been identified as a modulator for an amino acid sequence involved in the elongation of fatty acids (e.g. by means of a screen or assay as described hereinabove), it may be used per se as a modulator of the amino relevant amino acid sequence involved in the elongation of fatty acids (e.g. as an active substance for pharmaceutical use), or it may optionally be further optimized for final use, e.g. to improve properties such as solubility, ADME-TOX and other desired properties. It will be clear to the skilled person that the nucleotide sequences, amino acid sequences, host cells/host organisms and/or methods used in the invention may find further use in such optimization methodology, for example as (part of) secondary assays.

The invention is not particularly limited to any specific manner or mechanism in/via which the modulator (e.g. the test chemical, compound and/or factor) modulates, or interacts with, the target (*in vivo* and/or *in vitro*). For example, the modulator may an agonist, an antagonist, an inverse agonist, a partial agonist, a competitive inhibitor, a non-competitive inhibitor, a cofactor, an allosteric inhibitor or other allosteric factor for the target, and/or may be a compound or factor that enhances or reduces binding of target to another biological component associated with its (biological) activity, such as another protein or polypeptide, a receptor, or a part of organelle of a cell. As such, the modulator may bind with the target (at the active site, at an allosteric site, at a binding domain and/or at another site on the target, e.g. covalently or via hydrogen bonding), block the active site of the target (in a reversible, irreversible or competitive manner), block a binding domain of the target (in a reversible, irreversible or competitive manner), and/or influence or change the conformation of the target.

As such, the test chemical/modulator may for instance be:

- an analog of a known substrate of the target;
- an oligopeptide, e.g. comprising between 2 and 20, preferably between 3 and 15 amino acid residues;
- an antisense or double stranded RNA molecule;
- a protein, polypeptide;
- a cofactor or an analog of a cofactor.

Preferably, the compound is an inhibitor of the target, although the invention in its broadest sense is not limited thereto.

The test chemical/modulator may also be a reference compound or factor, which may be a compound that is known to modulate or otherwise interact with the target (e.g. a known substrate or inhibitor for the target) or a compound or factor that is generally known compound that is known to modulate or otherwise interact with other members from the general class to which the target belongs (e.g. a known substrate or inhibitor of said class).

Preferably, however, the compound(s) will be "small molecules", by which is generally meant herein a molecular entity with a molecular weight of less than 1500, preferably less than 1000. This may for example be an organic, inorganic or organometallic molecule, which may also be in the form or a suitable salt, such as a water-soluble salt; and may also be a complex, chelate and/or a similar molecular entities, as long as its (overall) molecular weight is within the range indicated above.

In a preferred embodiment, such a "small molecule" has been designed according, and/or meets the criteria of, at least one, preferably at least any two, more preferably at least any three, and up to all of the so-called Lipinski rules for drug likeness prediction (vide Lipinksi et al., Advanced Drug Delivery Reviews 23 (1997), pages 3-25). As is known in the art, small molecules which meet these criteria are particularly suited (as starting points) for the design and/or development of pharmaceuticals for human use, and may for instance be used as starting points for hits-to-leads chemistry, and/or as starting points for lead development (in which the methods used in the invention may also be applied).

Also, for these purposes, the design of such small molecules (as well as the design of libraries consisting of such small molecules) will preferably also take into account the presence of pharmacophore points, for example according to the methods described by I. Muegge et al., J. Med. Chem. 44, 12 (2001), pages 1-6 and the documents cited herein.

The term "small peptide " generally covers (oligo)peptides that contain a total of between 2 and 35, such as for example between 3 and 25, amino acids (e.g. in one or more connected chains, and preferably a single chain). It will be clear that some of these small peptides will also be included in the term small molecule as used herein, depending on their molecular weight.

In one preferred, but non-limiting embodiment, the invention is used to screen a set or library of (related or otherwise unrelated) small molecules, for example a standard "robustness set", a primary screening library (e.g. of otherwise unrelated compounds), a combinatorial library, a series of closely related chemical analogos. Such sets or libraries will be clear to the skilled person, and may for instance include, but are not limited to, such commercially available chemical libraries such as the various libraries available from Tocris Cookson, Bristol, UK.

The modulators thus identified by the methods used in the invention can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases, and/or can be used to develop other compounds that can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases, i.e. as already outlined above.

Also, as already mentioned above, the use of the human amino acid sequences of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33, and/or of sequences derived therefrom (such as analogs, parts, fragments, and/or fusions thereof as described hereinabove), and of host cells/host organisms containing/expressing these, are usually preferred, in particular when the invention is used to develop compounds for pharmaceutical use.

As already mentioned above, the compounds and/or factors that have been identified and/or developed as modulators of the amino acid sequences involved in the elongation of fatty acids (and/or precursors for such compounds) may be useful as active substances in the pharmaceutical field, for example in the preparation of pharmaceutical compositions, and both such modulators as well as (pharmaceutical) compositions containing them further aspects of the invention.

In particular, the compounds and composition of the invention may be used in (the preparation of pharmaceutical compositions for) the prevention (e.g. prophylaxis) and/or treatment of metabolic diseases (which for the purposes herein in its broadest sense also includes preventing, treating and/or alleviating the symptoms and/or complications of such metabolic diseases).

In particular, such compounds and composition may be used in (the preparation of pharmaceutical compositions for) the prevention (e.g. prophylaxis) and/or treatment of

metabolic diseases (which for the purposes herein in its broadest sense also includes preventing, treating and/or alleviating the symptoms and/or complications of such metabolic diseases).

In particular, the compounds and compositions of the invention may be used for preventing and/or treating:

- hyperglycemic conditions and/or other conditions and/or diseases that are (primarily) associated with (the response or sensitivity to) insulin, including but not limited to all forms of diabetes and disorders resulting from insulin resistance, such as Type I and Type II diabetes, as well as severe insulin resistance, hyperinsulinemia, and hyperlipidemia, e.g., obese subjects, and insulin-resistant diabetes, such as Mendenhall's Syndrome, Werner Syndrome, leprechaunism, lipoatrophic diabetes, and other lipoatrophies;
- conditions caused or usually associated with hyperglycemic conditions and/or obesity, such as hypertension, osteoporosis and/or lipodystrophy.
- so-called "metabolic syndrome" (also known as "Syndrome X") which is a condition where several of the following conditions coexist: hypertension; insulin resistance; diabetes; dyslipidemia; and/or obesity.

In particular, the compounds and compositions of the invention may be used for preventing and/or treating diabetes, especially Type I and Type II diabetes. "Diabetes" itself refers to a progressive disease of carbohydrate metabolism involving inadequate production or utilization of insulin and is characterized by hyperglycemia and glycosuria.

Also, as mentioned above, the amino acid sequences involved in the elongation of fatty acids and in particular the nucleotide sequences used in the invention, and more in particular the human amino acid sequences and nucleotide sequences used in the invention may be used for diagnostic purposes, for example as part of diagnostic assays and/or as part of kits for performing such assays (in which such a kit will comprise at least a nucleotide sequence used in the invention, may be suitably packaged (e.g. in a suitbale container) and may optionally further comprise one or more elements for such kits known per se, such as suitable reagents, buffers or other solvents, and instructions for use).

In particular, the amino acid sequences and nucleotide sequences used in the invention, as well as assays and kits using such sequences, may be used for diagnostic purposes relating to one or more of the metabolic diseases indicated above, for example as assays to determine the presense and/or absence in an individual of specific mutations and/or genetic markers that relate to one or more of the metabolic diseases referred to above, to determine the susceptibility and/or any predisposition for any of the metabolic diseases referred to above in an individual, to determine if any genetically determined factors contribute or even cause (in full or in part) a metabolic disease in an individual, determine and/or to confirm the kind of metabolic disease from which an individual suffers, and/or to predict the further progress of a metabolic disease in an individual. It will also be clear that any results obtained using such a diagnostic method or assay may also provide guidance to the clinician as to how a metabolic diseaase should be treated in an individual, e.g. which diet should be followed and/or which medication should be prescribed and/or the dosis regimen to be used.

It should also be noted that, for the treatment of the metabolic disease in humans, the compound used will usually and preferably be an inhibitor of an amino acid sequence involved in the elongation of fatty acids, although the invention is its broadest sense is not limited thereto.

In one specific, but non-limiting, embodiment of the invention, a compound is considered an inhibitor of one of the amino acid sequences involved in the elongation of fatty acids if, in a relevant assay such as the kinase activity assays referred to above (or a suitable modification thereof, for example using partially or fully purified protein), said compound reduces the activity of said amino acid sequence, i.e. by at least 1%, preferably at least 10%, such as by 20% or more, compared to the activity without the presence of said compound.

In an even more specific, but non-limiting, embodiment of the invention, a compound is considered an inhibitor of one of the amino acid sequences involved in the elongation of fatty acids if, in a relevant assay, such as a binding assay, said compound has an IC50 value of less than 1000 μm, preferably at than 500 μm, more preferably less than 250μM, even more preferably less than 100 μm, for example 50 μm or less, such as about 10 μm or less.

Again, preferably, in the invention compounds are used that are modulators, and in particular inhibitors, of the human amino acid sequences of SEQ ID NO: 3, 5, 8, 11, 14, 17, 20, 23, 26 and 33, and/or of amino acid sequences derived therefrom, such as analogs, mutants, parts, fragments and/or fusions as described above.

For pharmaceutical use, the compounds of the invention may be used as a free acid or base, and/or in the form of a pharmaceutically acceptable acid-addition and/or base-addition salt (e.g. obtained with non-toxic organic or inorganic acid or base), in the form of a hydrate, solvate and/or complex, and/or in the form or a pre-drug, such as an ester. Such salts, hydrates, solvates, etc. and the preparation thereof will be clear to the skilled person; reference is for instance made to the salts, hydrates, solvates, etc. described in US-A-6,372,778, US-A-6,369,086 and US-6,369,067

Generally, for pharmaceutical use, the compounds of the inventions may be formulated as a pharmaceutical preparation comprising at least one compound of the invention and at least one pharmaceutically acceptable carrier, diluent or excipient and/or adjuvant, and optionally one or more further pharmaceutically active compounds. By means of non-limiting examples, such a formulation may be in a form suitable for oral administration, for parenteral administration (such as by intravenous, intramuscular or subcutaneous injection or intravenous infusion), for topical administration, for administration by inhalation, by a skin patch, by an implant, by a suppository, etc.. Such suitable administration forms - which may be solid, semi-solid or liquid, depending on the manner of administration - as well as methods and carriers for use in the preparation thereof, will be clear to the skilled person; reference is again made to for instance US-A-6,372,778, US-A-3,696, 086 and US-6,369,067.

The pharmaceutical preparations of the invention are preferably in a unit dosage form, and may be suitably packaged, for example in a box, blister, vial, bottle, sachet, ampoule or in any other suitable holder or container (which may be properly labeled); optionally with one or more leaflets containing product information and/or instructions for use. Generally, such unit dosages will contain between 1 and 500 mg of the at least one compound of the invention, e.g. about 10, 25, 50, 100, 200, 500 or 1000 mg per unit dosage.

For pharmaceutical use, at least one compound of the invention will generally be administered in an amount of between 0.01 to 150 mg/kg body weight per day of the patient, divided over one or more daily doses. The amount(s) to be administered and the further treatment regimen may be determined by the treating clinician, depending on factors such as the age, gender and general condition of the patient and the nature and severity of the disease/symptoms to be treated.

Thus, in a further aspect, the invention relates to a composition, and in particular a composition for pharmaceutical use, that contains at least one compound of the invention (i.e. a compound that has been identified, discovered and/or developed using a nematode or method as described herein) and at least one suitable carrier (i.e. a carrier suitable for pharmaceutical use). The invention also relates to the use of a compound of the invention in the preparation of such a composition.

Preferably, the compounds and compositions of the invention are administered orally and/or in a form intended and/or suitable for oral administration.

It is also envisaged that the above compounds and compositions may be of value in the veterinary field, which for the purposes herein not only includes the prevention and/or treatment of diseases in animals, but also - for economically important animals such as cattle, pigs, sheep, chicken, fish, etc. - enhancing the growth and/or weight of the animal and/or the amount and/or the quality of the meat or other products obtained from the animal. Thus, in a further aspect, the invention relates to a composition for veterinary use that contains at least one compound of the invention (i.e. a compound that has been identified, discovered and/or developed using a nematode or method as described herein) and at least one suitable carrier (i.e. a carrier suitable for veterinary use). The invention also relates to the use of a compound of the invention in the preparation of such a composition.

In the agrochemical field, the invention may be used to identify compounds suitable for use in pesticides, insecticides, nematicides and/or other biocides or plant protection agents. For example, the compounds invention may be used to control the species listed in US-A-6,372,774. For this purpose, the compounds of the invention (or a suitable salt, hydrate or ester thereof) may be suitably formulated with one or more agrochemically acceptable carriers, to provide a formulation suitable for agrochemical

use, as will be clear to the skilled person (reference is for example made to the formulations and uses described in US-A-6,372,774).

Thus, in a further aspect, the invention relates to a composition for agrochemical use that contains at least one compound of the invention (i.e. a compound that has been identified, discovered and/or developed using a nematode or method as described herein) and at least one suitable carrier (i.e. a carrier suitable for agrochemical use). The invention also relates to the use of a compound of the invention in the preparation of such a composition.

The invention will now be further illustrated by means of the following non-limiting Experimental Part.

In the Figures:

- Figure 1 schematically shows vector pGN49A (see also WO 00/01846 and UK patent application no. GB 0012233, both by Applicant);
- Figures 2A and 2B are photographs (enhanced using the Scion Image (Scion Corp) software package) showing reduced fat-absorption phenotype in *C. elegans* upon Nile Red Staining: Figure 2A = reduced fat storage (invention); Figure 2B = reference (reference vector gGN29 without RNAi insert).

Experimental part:

In the Experimental Part below, unless indicated otherwise, all steps for handling and cultivating *C. elegans* were performed using standard techniques and procedures, for which reference is made to the standard *C. elegans* handbooks, such as W.B. Wood et al., "*The nematode Caenorhabditis elegans*", Cold Spring Harbor Laboratory Press (1988); D.L. Riddle et al., "*C. ELEGANS II*", Cold Spring Harbor Laboratory Press (1997); "*Caenorhabditis elegans, Modern Biological analysis of an organism*": ed. by H. Epstein and D. Shakes, Methods in Cell Biology, Vol 48, 1995; and "*C. elegans, a practical approach*", ed. by I.A. Hope, Oxford University Press Inc. New York, USA, 1999.

Downregulation of the gene(s) of interest in *C. elegans* was achieved by RNAi feeding techniques using an *E.coli* strain capable of expressing a dsRNA corresponding to the gene(s) of interest, as generally described in - *inter alia* - the International application WO 00/01846 by applicant and the handbooks referred to above.

Also, unless indicated otherwise, all cloning and other molecular biology steps were performed using standard techniques and protocols, i.e. as provided by the manufacturers of the reagents/kits used and/or as described in the standard handbooks, such as Sambrook et al, "Molecular Cloning: A Laboratory Manual" (2nd.ed.), Vols. 1-3, Cold Spring Harbor Laboratory Press (1989) and F. Ausubel et al, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987).

Fat accumulation in *C. elegans daf-2 (e1370)* was determined visually under a microscope upon staining with Nile-red, using an adaptation of the general methodology described by Ogg et al., Nature, Vol. 389, 994 (1997). For the general methodology, reference is also made to Thaden et al., 1999 International Worm Meeting abstract 837; Ashrafi and Ruvkun, 2000 East Coast Worm Meeting abstract 67; Ashrafi, Chang and Ruvkun, 2001 International Worm Meeting abstract 325; and Rottiers and Antebi, 2001 International Worm Meeting abstract 620 (all abstracts available from Worm Literature Index at http://elegans.swmed.edu/wli/).

Example 1: Preparation *E.coli* RNA feeding strain for expression of C56G2.6 double stranded RNA.

A vector for expression of dsRNA for downregulation of *C. elegans* gene C56G2.6 was prepared as follows.

The DNA fragment of SEQ ID NO: 30, which corresponds to 821 nucleotides of the *C. elegans* C56G2.6 gene (SEQ ID NO.27), was obtained by PCR from genomic *C. elegans* DNA, using the following primers:

- forward primer : TGCCAGTGCT TCTTGGTTGG [SEQ ID NO: 28]
- reverse primer : AGCTCGAGTT GAAGTTGATG CG [SEQ ID NO: 29]

This fragment was inserted in the *SrfI*-site of expression vector pGN49a pGN49A (Figure 1, see also WO 00/01846 and UK patent application no. GB 0012233, both by Applicant). This vector contains two T7 promoters flanking the *SrfI*-site, allowing

transcription of a nucleotide sequence inserted into said *SrfI*-site into double stranded RNA, upon binding of a T7 polymerase to said promoter (vide WO 00/01846).

The resulting vector, designated pGN49A- C56G2.6, was transformed overnight into *E. coli* strain AB 309-105 (see EP-A-1 093 526 by applicant, page 17.).

To normalize the culture, 250 µl of the overnight culture (1 ml) was transferred to a 96 well plate and the OD at 600 nm was measured (Fluostar Galaxy plate reader BMG), the remaining 750 µl centrifuged down. Next the pellet was re-suspended in S-complete fed (S-complete supplemented with 0.1mg/ml ampiciline and 1 mM IPTG) and volume adjusted to obtain OD_{600} value of 1

Example 2: Generation of fat storage phenotype in *C. elegans* - P0 screen for *C. elegans* gene C56G2.6.

In this example, *C. elegans* strain CB1370 containing the temperature sensitive daf-2 allele e-1370 is used (Ogg et al., *supra*). CB 1370 is publicly available from, for example, the Caenorhabditis Genetics Center (CGC), Minnesota, USA).

To generate the fat-storage phenotype, L1 worms of strain CB 1370 were cultivated at a temperature of 15 °C in S-Complete fed-medium in the wells of a 96 well plate (40 L1 nematodes per well) under essentially synchronized conditions, until the nematodes reached the L2 stage.

Then, the temperature was increased to 25°C, and the worms were further cultivated at said temperature until they reached the L4 stage (about 36-48 hours). Due to the presence of the daf-2 allele e-1370, this raise in temperature from 15°C to 25°C causes the nematodes to accumulate fat, mainly in their intestinal and hypodermal tissue (vide Ogg et al. and Figures 2A and 2B).

The accumulation of fat (in the form of droplets) was made visible by means of Nile Red staining: L4 animals were washed several times with M9 (supplemented with 0.1% PEG) to remove the remaining *E.coli*, and fixed with MeOH (fc. 33%). After fixation the nematodes were stained with nile red (fc 0.375 mM in 37.5% MeOH) for 4 hours. MeOH and excess dye was removed through several washes with M9 (supplemented with 0.1% PEG). The staining pattern was visualized under UV using a 500 nm long pass filter.

For testing the influence of the gene C56G2.6 on fat storage, during the steps described above, the worms were grown on 20 μl of the normalized *E.coli* strain containing the pGN49A vector with the RNAi fragment for C56G2.6 inserted therein, as obtained in Example 1 (OD_{600} = 1), as a food source. As a reference, the daf-2 (e1370) nematodes were grown in a similar manner, but with *E. coli* strain AB 309-105 containing vector pGN49A without the RNAi fragment for C56G2.6 inserted therein as a food source, used in the same amount. All samples were carried out in quadruplicate.

The results were as follows: worms fed on *E.coli* pGN49A- C56G2.6 strain, which downregulates the expression of C56G2.6 through RNA interference, showed a strong reduction of the accumulation of fat, compared to the reference (vide Figures 2A and 2B).

These results show that C56G2.6 is involved in the regulation of (the daf-2 dependent) accumulation of fat in the nematode. From bio-informatic analysis, it was found that Genbank protein 7705855/Genbank DNA HSD17B12 is the closest human ortholog for C56G2.6. As already mentioned above, HSD17B12 (nucleotide sequence NM_016142; amino acid sequence NP057226.1) corresponds to KAR mentioned above. This result confirms that enzymes that are involved in the elongation of fatty acids (as described above) can be used as targets in metabolic disease.

Also, besides KAR, from bio-informatic analysis, a close human homolog of both KAR and the *C. elegans* gene C56G2.6 was also established, the nucleotide sequence of which is given in SEQ ID NO: 31 (mRNA sequence) and SEQ ID NO: 32 (mRNA - coding sequence); the amino acid sequence is given in SEQ ID NO: 33. Reference is also made to Genbank accession numbers NM_031463 (gi 22432036) for the nucleotide sequence and NP_113651.3 (gi 22432037) for the amino acid sequence. Without being limited to any specific hypothesis or explanation, it is assumed that this specific amino acid sequence is also involved in the elongation of fatty acids, and possibly in the Second Step referred to above. Therefore, the use of the sequences of SEQ ID NOS 31, 32 and in particular 33 (as well as anologs thereof, as defined hereinabove) in the methods described herein form a further aspect of the invention.

CLAIMS:

1. Method for identifying a compound that can be used in the prevention and/or treatment of metabolic diseases, said method at least comprising the steps of:

a) contacting an amino acid sequence that is involved in the elongation of fatty acids, and/or a host cell or host organism containing/expressing such an amino acid sequence, with a test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said amino acid sequence; and optionally

b) detecting the signal that may thus be generated, said signal identifying a modulator of said amino acid sequence;

in which the modulator thus identified can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases.

2. Method for generating a signal that is representative for the interaction of a test chemical with amino acid sequence involved in the elongation of fatty acids, said method at least comprising the steps of:

a) contacting said amino acid sequence, or a host cell or host organism containing/expressing said amino acid sequence involved in the elongation of fatty acids, with said test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said amino acid sequence; and optionally

b) detecting the signal that may thus be generated.

3. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

4. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity for fatty acids with a length of the carbon chain of 20 carbon atoms or less; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

5. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity for fatty acids with a length of the carbon chain of 18 carbon atoms or less; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

6. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity for saturated fatty acids with 18 carbon atoms or less in the fatty acid chain, compared to both (1) saturated and unsaturated fatty acids with more than 18 carbon atoms in the fatty acid chain; as well as (2) unsaturated fatty acids of the (n-3) family and/or (n-6) family with 18 carbon atoms or less in the fatty acid chain; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

7. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity for palmitic acid and/or for stearic acid, compared to both (1) saturated and unsaturated fatty acids with more than 18 carbon atoms in the fatty acid chain; as well as (2) linoleic acid and/or alpha-linolenic acid; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

8. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of:

- condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA that have specificity (as determined by a suitable assay, such as one of the assays mentioned hereinabove) for palmitic acid, compared to both (1) stearic acid; as well as (2) linoleic acid and alpha-linolenic acid; and
- reductases involved in either the reduction of the 3-ketoacyl-CoA using NADPH to form 3-hydroxyacyl-CoA; and/or in the reduction of trans-2,3-enoyl-CoA to saturated acyl-CoA.

9. Method according to any of the preceding claims, in which the condensation enzymes involved in the condensation between a fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA contains at least a histidin-rich motif (HXXHH), and prefereably also contains an ELO-domain.

10. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is derived from a mammal, and in particular from a human, or is an analog of such an amino acid sequence derived from a mammal, and in particular a human.

11. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of: Cig30, Ssc1, Ssc2, LCE, ELOVL1, ELOVL2, ELOVL3, ELOVL4, ELOVL5 (HELO1), ELOVL6, KAR and TER, and natural or synthetic analogs thereof, and in particular from the group consisting of: LCE, ELOVL6, KAR and TER, and natural or synthetic analogs thereof.

12. Method according to any of the preceding claims, in which the amino acid sequence involved in elongation of fatty acids is chosen from the group consisting of SEQ ID NOS: 3, 5, 8, 11, 14, 17, 20, 23, 26 and/or 33, and natural or synthetic analogs thereof, an in particular from the group consisting of: SEQ ID NO: 3, 23 and/or 26, and natural or synthetic analogs thereof.

13. Compound that can be used in the treatment of metabolic diseases, identified by a method of any of claims 1-12.

14. Compound according to claim 13, in which said compound is a modulator of an amino acid sequence that is involved in the elongation of fatty acids

15. Modulator of an amino acid sequence that is involved in the elongation of fatty acids, identified by a method of any of claims 1-12.

16. Pharmaceutical composition, comprising at least one compound or modulator according to any of claims 11-13 and at least one pharmaceutically acceptable carrier.

17. Pharmaceutical composition according to claim 13, being a composition suitable and/or intended for oral administration.

18. Use of a compound or modulator according to any of claims 11-13 in the preparation of a pharmaceutical composition.

19. Use of a compound or modulator according to any of claims 11-13 in the reparation of a pharmaceutical composition for the prevention and/or treatment of metabolic diseases.

Figure 1

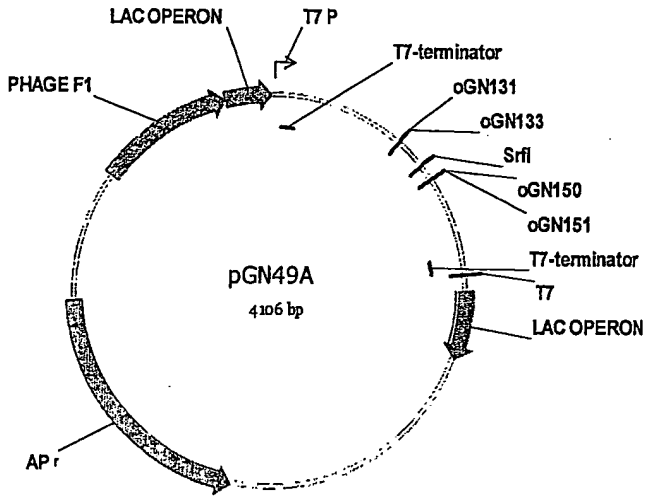

LAC OPERON
T7 P
T7-terminator
PHAGE F1
oGN131
oGN133
SrfI
oGN150
oGN151
T7-terminator
T7
LAC OPERON
pGN49A
4106 bp
AP r

Figure 2A

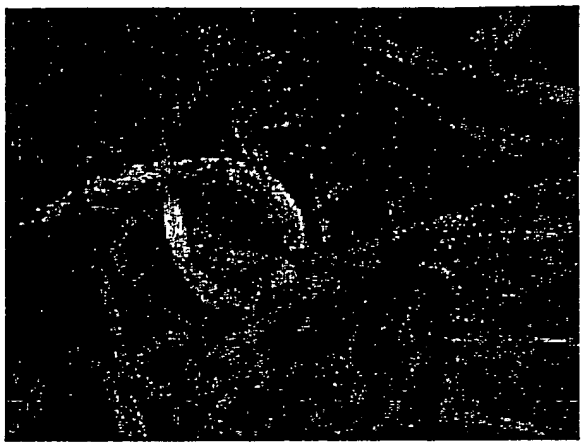
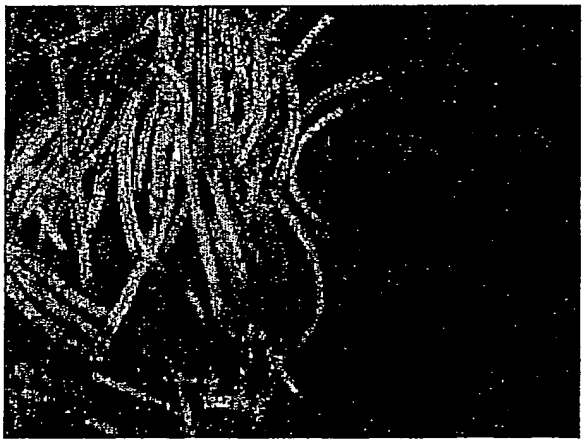


Figure 2B



SEQUENCE LISTING

<110> Devgen N.V.

<120> Use of amino acid sequences involved in the elongation of fatty acids in identifying and/or developing compounds for preventing and/or treating of metabolic diseases.

<130> P03-015

<160> 33

<170> PatentIn version 3.1

<210> 1

<211> 3045

<212> DNA

<213> homo sapiens

<400> 1
actaagaccg caaggcattc atttcctcct acggtggatg cggacgccgg gaggaggaga 60

gccccagaga gaggagctgg gagcggaggc gcaggcaatg ctcagccctg gatgtagctg 120

agaggctggg agaagagacg accgctggag accgagcggc gtggggaaga cctaggggg 180

tgggtggggg aagcagacag gagaacactc gaaatcaagc gctttacaga ttattttatt 240

ttgtatagag aacacgtagc gactccgaag atcagcccca atgaacatgt cagtgttgac 300

tttacaagaa tatgaattcg aaaagcagtt caacgagaat gaagccatcc aatggatgca 360

ggaaaactgg aagaaatctt tcctgttttc tgctctgtat gctgccttta tattcggtgg
420

tcggcaccta atgaataaac gagcaaagtt tgaactgagg aagccattag tgctctggtc
480

tctgaccctt gcagtcttca gtatattcgg tgctcttcga actggtgctt atatggtgta
540

catttgatg accaaaggcc tgaagcagtc agtttgtgac cagggttttt acaatggacc
600

tgtcagcaaa ttctgggctt atgcatttgt gctaagcaaa gcacccgaac taggagatac
660

aatattcatt attctgagga agcagaagct gatcttcctg cactggtatc accacatcac
720

tgtgctcctg tactcttggt actcctacaa agacatggtt gccgggggag gttggttcat
780

gactatgaac tatggcgtgc acgccgtgat gtactcttac tatgccttgc gggcggcagg
840

tttccgagtc tcccggaagt ttgccatgtt catcaccttg tcccagatca ctcagatgct
900

gatgggctgt gtggttaact acctggtctt ctgctggatg cagcatgacc agtgtcactc
960

tcactttcag aacatcttct ggtcctcact catgtacctc agctaccttg tgctcttctg
1020

ccatttcttc tttgaggcct acatcggcaa aatgaggaaa acaacgaaag ctgaatagtg
1080

ttggaactga ggaggaagcc atagctcagg gtcatcaaga aaaataatag acaaaagaaa
1140

atggcacaag gaatcacacg tggtgcagct aaaacaaaac aaaacatgag caaacacaaa
1200

acccaaggca gcttagggat aattaggttg atttaaccca gtaagtttat gatcctttta
1260

gggtgaggac tcactgagtg cacctccatc tccaagcact gctgctggaa gaccccattc
1320

cctctttatc tatcaactct aggacaaggg agaacaaaag caagccagaa gcagaggaga
1380

ctaatcaaag gcaaacaaag gctattaaca cataggaaaa tatgtattta ctaagtgtca
1440

catttctcta agatgaaaga tttttactct agaaactgtg cgagcacaac acacacaatc
1500

ctttctaact ttatggacac taaactggag ccaatagaaa agacaaaaat gaaagagaca
1560

cagggtgtat atctagaacg ataatgcttt tgcagaaact aaagcctttt taagaaatgc
1620

cagctgctgt agaccccatg agaaaagatg tcttaatcat ccttatgaaa acagatgtaa
1680

acaactatat ttcaactaac ttcatcttca ctgcatagcc tcaggctagt gagtttgcca
1740

aaaccaaagg gggtgaatac ttccccaaga ttcttcctgg gaggatggaa acagtgcagc
1800

ccaggtccca tgggggcagc tccatcccag agcatttctg atagttgaac tgtaatttct
1860

actcttaagt gagatatgaa gtattatcct tttgttcagt tgccccgggc ttttgaacag
1920

aagagtaaat acagaattga aaaagataaa cactcaacca aacaatgtga aaacgggttc
1980

tgtagtattt gtaaaaaggc ccggcccagg accactgtga gctggaaaag ggagaaaggc
2040

agtgggaaaa gaggtgagcc gaagatcaat tcgacagaca gacggtgtgt atgcccctcc
2100

ctgtttgact tcacacacac tcataacttt ccaaatgaaa ccccacagta tagcgcatat
2160

tttcgatatt tttgtgaatt ccaaaaggaa atcacagggc tgttcgaaat attgggggaa
2220

cactgtgttt ctgcatcatc tgcatttgct ccccaagcaa tgtagaggtg tttaaagggc
2280

cctctgctgg ctgagtggca atactacaac aaacttcaag gcaagtttgg ctgaaaacag
2340

ttgacaacaa agggcccccа tacacttatc cctcaaattt taagtgatat gaaatacttg
2400

tcatgtcttt ggccaaatca gaagatattc atcctgcttc aagtcagctt cagaaatgtt
2460

ttaaaaggga ctttagctct ggaactcaaa atcaatttat taagagccat attctttaaa
2520

aaaaaaagct ggataatatt atctgtaata tttcagtcct ttacaagcca aatacatgtg
2580

tcaatgtttc tagtatttca aagaagcaat tatgtaaagt tgttcaatgt gacataatag
2640

tattataatt ggttaagtag cttaatgatt aggcaaacta gatgaaaaga ttaggggctt
2700

ccacactgca tagatcacac gcacatagcc acgcatacac acacagacac acagatgtgg
2760

ggtacactga atttcaaagc ccaaatgaat agaaacacat tttctggcta gcagaaaaaa
2820

acaaaacaaa actgttgttt ctctttcttg ctttgagagt gtacagtaaa agggattttt
2880

tcgaattatt tttatattat tttagcttta attgtgctgt cgttcatgaa acagagctgc
2940

tctgcttttc tgtcagagat ggcaagggct ttttcagcat ctcgtttatg tgtggaattt
3000

aaaaagaata aagttttatt ccattctgaa aaaaaaaaaa aaaaa
3045

<210> 2

<211> 798

<212> DNA

<213> Homo sapiens

<400> 2
atgaacatgt cagtgttgac tttacaagaa tatgaattcg aaaagcagtt caacgagaat
60

gaagccatcc aatggatgca ggaaaactgg aagaaatctt tcctgttttc tgctctgtat
120

gctgccttta tattcggtgg tcggcaccta atgaataaac gagcaaagtt tgaactgagg
180

aagccattag tgctctggtc tctgaccctt gcagtcttca gtatattcgg tgctcttcga
240

actggtgctt atatggtgta cattttgatg accaaaggcc tgaagcagtc agtttgtgac
300

cagggttttt acaatggacc tgtcagcaaa ttctgggctt atgcatttgt gctaagcaaa
360

gcacccgaac taggagatac aatattcatt attctgagga agcagaagct gatcttcctg
420

cactggtatc accacatcac tgtgctcctg tactcttggt actcctacaa agacatggtt
480

gccgggggag gttggttcat gactatgaac tatggcgtgc acgccgtgat gtactcttac
540

tatgccttgc gggcggcagg tttccgagtc tcccggaagt ttgccatgtt catcaccttg
600

tcccagatca ctcagatgct gatgggctgt gtggttaact acctggtctt ctgctggatg
660

cagcatgacc agtgtcactc tcactttcag aacatcttct ggtcctcact catgtacctc
720

agctaccttg tgctcttctg ccatttcttc tttgaggcct acatcggcaa aatgaggaaa
780

acaacgaaag ctgaatag
798

<210> 3

<211> 265

<212> PRT

<213> Homo sapiens

<400> 3

Met Asn Met Ser Val Leu Thr Leu Gln Glu Tyr Glu Phe Glu Lys Gln
1 5 10 15

Phe Asn Glu Asn Glu Ala Ile Gln Trp Met Gln Glu Asn Trp Lys Lys
20 25 30

Ser Phe Leu Phe Ser Ala Leu Tyr Ala Ala Phe Ile Phe Gly Gly Arg
35 40 45

His Leu Met Asn Lys Arg Ala Lys Phe Glu Leu Arg Lys Pro Leu Val
50 55 60

Leu Trp Ser Leu Thr Leu Ala Val Phe Ser Ile Phe Gly Ala Leu Arg
65 70 75 80

Thr Gly Ala Tyr Met Val Tyr Ile Leu Met Thr Lys Gly Leu Lys Gln
85 90 95

Ser Val Cys Asp Gln Gly Phe Tyr Asn Gly Pro Val Ser Lys Phe Trp
100 105 110

Ala Tyr Ala Phe Val Leu Ser Lys Ala Pro Glu Leu Gly Asp Thr Ile
115 120 125

Phe Ile Ile Leu Arg Lys Gln Lys Leu Ile Phe Leu His Trp Tyr His
130 135 140

His Ile Thr Val Leu Leu Tyr Ser Trp Tyr Ser Tyr Lys Asp Met Val
145 150 155 160

Ala Gly Gly Gly Trp Phe Met Thr Met Asn Tyr Gly Val His Ala Val
165 170 175

Met Tyr Ser Tyr Tyr Ala Leu Arg Ala Ala Gly Phe Arg Val Ser Arg
180 185 190

Lys Phe Ala Met Phe Ile Thr Leu Ser Gln Ile Thr Gln Met Leu Met
195 200 205

Gly Cys Val Val Asn Tyr Leu Val Phe Cys Trp Met Gln His Asp Gln
210 215 220

Cys His Ser His Phe Gln Asn Ile Phe Trp Ser Ser Leu Met Tyr Leu
225 230 235 240

Ser Tyr Leu Val Leu Phe Cys His Phe Phe Phe Glu Ala Tyr Ile Gly
245 250 255

Lys Met Arg Lys Thr Thr Lys Ala Glu
260 265

<210> 4

<211> 3324

<212> DNA

<213> Homo sapiens

<400> 4
aaatttattg agctattaga tacgatcttt tttgttctgc gcaagaaaaa tagccaagtg
60

actttccttc atgtattcca tcataccatc atgccgtgga cctggtggtt tggagtcaaa
120

tttgctgcag gtggtttggg aacattccat gcccttctaa atacagctgt acatgtagtc
180

atgtattcct actatggact ttctgcattg gggccagcct accagaagta tttgtggtgg
240

aaaaaatatt tgacatcatt acagcttgtc cagtttgtta ttgtcgccat ccacataagc
300

cagttctttt tcatggagga ttgcaagtat cagtttccag tctttgcgtg catcattatg
360

agttacagtt tcatgtttct gctgctcttt ctccattttt ggtaccgtgc ttacaccaaa
420

ggtcagaggt tgcccaaaac tgtgaaaaat ggaacttgca aaaacaaaga taattgaagc
480

ccaacataag tctatgatcg aaactgatac attgtcttcc ttgacaatca agagatattt
540

acctatgcag tgcattttgt atatttttca aaactaagag cttgtatttt tatggtaagt
600

tattgggatg tctgatattt gagagcccga agcttccaga taacagtctt ctgataatta
660

gagcctacag cagccagaat ttgtttttgt ttttgttttt gttttttag ctgcttttaa
720

acctaatcca aaggttttgt ttaaaacatt ttgttgcagt gaaaaaaaga tatgaagcaa
780

tacagtactc ttcaaagaag tccaataaag atgaaaaata tcattaggta ttttgagcac
840

agacggagat ccatgtgata aaatagattc cttctgctgg gtctggaaag tctggaaacc
900

acctgtaggc cctggctttg tcatttattg actgtattca gaagatacta tatttgctct
960

agagttaact cccattttag caagctagca caggtgaaaa ttgagtgagt ttttgataac
1020

ttgtcattta aaatcattaa tgataatttt caatggatct tttcagtagc catcaccagt
1080

tttgctgata agacttcctc acaacccatt tgttgtacaa actgtttcaa agtagcaatc
1140

ctttgggtta gtacacttga taccaagttt cattcagttg attattttca aaacaaggtg
1200

atttgtttta atgggttaat gaatactttg ctattactgt tttacaatta actttgtata
1260

tctctggaaa gaggtgaatt tgtcaatgaa aaaagtattg tgtagttcag tgggaaaaac
1320

ctgttgtctg ttatagtatc acatcacttt cttacattgt catggtttaa attattatct
1380

tggtgaaata tataggttag atacttaaat gttcatatta ttagcacact acaggtacta
1440

cccttaaatc ataaataata ttacccattg tacacagttt ccagatatga gttagagtgg
1500

ctgtgaagca actagaggca aattgtggca cagagaatat ccaggggaaa attgattatg
1560

taagcaaggg ctgttctact ttgagagaga gacagacaga ctacataata gtaatatatt
1620

atataattaa tgttatagat acactgaaca ataaaacatt ttctggaatt atgaattaat
1680

taacaacctg gatttctgtt tccagaccta taaaccatgt aatgaaggac tgaaattcgt
1740

ttggcataca ctttgtttct ttaaaattgc tagtttcttt ctgttatttt tacattttct
1800

tgtcagagaa tcaaaactat tagtcagtag agttttgtcc atgaaataat atttggacat
1860

ttgtgaattt ttccctattt tttttcttct tttatttata ctcaattttg aagcactgtt
1920

tatgtttgta ggactttaaa caattagtac ttaaagcccc agttaatttt gaacacacag
1980

agatacttat tgtactctgt gtacagtaag tatttttatg tttacactta ccagattcat
2040

aaggttattt gcctttaagt gatcttttgt gattttactt gattacagca tgagaaggta
2100

aaggttgata aatgggagta attacatata tatatataat ttattatttt tttttctgaa
2160

gcagagtttt actcttgttg cccaggctag agtgcaatgg cacaatctcg gttcactgca
2220

acctctgcct cccaggttca agtgattctc ctgtctcagc ctcccaagta gctgggatta
2280

caggtgcccg ccaccatgct cagctaattt ttgtattttt agtagagaca gggttggcca
2340

ggctggtctc gagctcctga cctcaggtga tcctcccacc tcggcctccc aaagttctgg
2400

gattacaggg gtgagccact gtgcccggcc ataattttat atttttagaa aataaagaat
2460

gaaggacaaa agagcacagc taaatgaaac tgctgctcta atttattcca ctggaaaggt
2520

ctcaggactc cttaactgtt ttccaggttt ggctcttcat atctaacctg tgctaaaatg
2580

agaaggtatg ctaggtgctg ttagaaatca caccatttcc aaagaaccca agtagtatag
2640

agaccaaaat gaggcaaaaa taaaagagat gaaacaggag agtttatttc ttgcatactt
2700

tccgaccttc atcatacaac ctctcctaac ctccctgtag tctttaaaat gtttaacttg
2760

cctataagct aacatgtaat aaaacacata ctcaattata tgaatagaag tggagagcca
2820

gaatgctaca aaagaaatca cagctgctag aagtatcctc catagaaaca tcaacatgat
2880

tgtggatcaa aatgattttc actggatatg gaatttgtat gggccatatt tattaaaaga
2940

gttctgtgtg gtcacataga gttcctttgg gatttcatcc cacttttcag gacttaattt
3000

gtttgggttt gcttacctaa caagcaacaa caacaacaaa ataaatagtt ccaaaatcta
3060

gtttattaca taaatctcta tgaatccttg agggcctata tctatattat aaataaatac
3120

aaatatagat ttttaaatat ctatgggact ttgccattta cagccttaag tataaaatta
3180

cgagattata ttctttccat tacccttttat ttctgctaac tttttaaaga ctggaacatt
3240

ctgtggatgt tgtcaaagtt tgagtttgtt tttccctgtg tattataata aatttgtggt
3300

attgcaaaaa aaaaaaaaaa aaaa
3324

<210> 5

<211> 158

<212> PRT

<213> Homo sapiens

<400> 5

Lys Phe Ile Glu Leu Leu Asp Thr Ile Phe Phe Val Leu Arg Lys Lys
1 5 10 15

Asn Ser Gln Val Thr Phe Leu His Val Phe His His Thr Ile Met Pro
20 25 30

Trp Thr Trp Trp Phe Gly Val Lys Phe Ala Ala Gly Gly Leu Gly Thr
35 40 45

Phe His Ala Leu Leu Asn Thr Ala Val His Val Val Met Tyr Ser Tyr
50 55 60

Tyr Gly Leu Ser Ala Leu Gly Pro Ala Tyr Gln Lys Tyr Leu Trp Trp
65 70 75 80

Lys Lys Tyr Leu Thr Ser Leu Gln Leu Val Gln Phe Val Ile Val Ala
85 90 95

Ile His Ile Ser Gln Phe Phe Phe Met Glu Asp Cys Lys Tyr Gln Phe
100 105 110

Pro Val Phe Ala Cys Ile Ile Met Ser Tyr Ser Phe Met Phe Leu Leu
115 120 125

Leu Phe Leu His Phe Trp Tyr Arg Ala Tyr Thr Lys Gly Gln Arg Leu
130 135 140

Pro Lys Thr Val Lys Asn Gly Thr Cys Lys Asn Lys Asp Asn
145 150 155

<210> 6

<211> 1347

<212> DNA

<213> Homo sapiens

<400> 6
cgccgccgcg cacatgccta ggttcgacgc cctcctccct ttgcccagga gttccttctg
60

tcccggctct gttccgtctc gccccgaggt tcacgccatc ctcggagccc cagcctttca
120

cccagcgcct ccaagctttg gaccttgact tctgcaaaac tagatggtca cagccatgaa
180

tgtctcacat gaagtaaatc agctgttcca gccctataac ttcgagctgt ccaaggacat
240

gaggcccttt ttcgaggagt attgggcaac ctcattcccc atagccctga tctacctggt
300

tctcatcgct gtggggcaga actacatgaa ggaacgcaag ggcttcaacc tgcaagggcc
360

tctcatcctc tggtccttct gccttgcaat cttcagtatc ctgggggcag tgaggatgtg
420

gggcattatg gggactgtgc tacttaccgg gggcctaaag caaaccgtgt gcttcatcaa
480

cttcatcgat aattccacag tcaaattctg gtcctgggtc tttcttctca gcaaggtcat
540

agaactcgga gacacagcct tcatcatcct gcgtaagcgg ccactcatct ttattcactg
600

gtaccaccac agcacagtgc tcgtgtacac aagctttgga tacaagaaca aagtgcctgc
660

aggaggctgg ttcgtcacca tgaactttgg tgttcatgcc atcatgtaca cctactacac
720

tctgaaggct gccaacgtga agcccccaa gatgctgccc atgctcatca ccagcctgca
780

gatcttgcag atgtttgtag gagccatcgt cagcatcctc acgtacatct ggaggcagga
840

tcagggatgc cacaccacga tggaacactt attctggtcc ttcatcttgt atatgaccta
900

tttcatcctc tttgcccact tcttctgcca gacctacatc aggcccaagg tcaaagccaa
960

gaccaagagc cagtgaaggt ttggagagaa caatgaagct ccaggctctc tcttctccag
1020

ggcaccaaga ggctgggctt agttttggga gaatgattag gttgccttac ctgcatggtt
1080

tccccagagg atgtgtgccc caaggtggct ggaatttttg acagacaaga agggtgacct
1140

tgggatgggg gtgtggtctg ttactttaat gtttctgttt ttaatgtgaa ggccaagcag
1200

gccctgggat gggagtgggg cggaggaggg tcctaagagc tgattattta atttctatcc
1260

agaaatcttt cttcttcttg ctctgttttt ttaaattaaa gatttcaaca aaaaaaaaaa
1320

aaaaaaaaaa aaaaaaaaaa aaaaaaa
1347

<210> 7

<211> 813

<212> DNA

<213> Homo sapiens

<400> 7
atggtcacag ccatgaatgt ctcacatgaa gtaaatcagc tgttccagcc ctataacttc
60

gagctgtcca aggacatgag gccctttttc gaggagtatt gggcaacctc attccccata
120

gccctgatct acctggttct catcgctgtg gggcagaact acatgaagga acgcaagggc
180

ttcaacctgc aagggcctct catcctctgg tccttctgcc ttgcaatctt cagtatcctg
240

ggggcagtga ggatgtgggg cattatgggg actgtgctac ttaccggggg cctaaagcaa
300

accgtgtgct tcatcaactt catcgataat tccacagtca aattctggtc ctgggtcttt
360

cttctcagca aggtcataga actcggagac acagccttca tcatcctgcg taagcggcca
420

ctcatcttta ttcactggta ccaccacagc acagtgctcg tgtacacaag ctttggatac
480

aagaacaaag tgcctgcagg aggctggttc gtcaccatga actttggtgt tcatgccatc
540

atgtacacct actacactct gaaggctgcc aacgtgaagc cccccaagat gctgcccatg
600

ctcatcacca gcctgcagat cttgcagatg tttgtaggag ccatcgtcag catcctcacg
660

tacatctgga ggcaggatca gggatgccac accacgatgg aacacttatt ctggtccttc
720

atcttgtata tgacctattt catcctcttt gcccacttct tctgccagac ctacatcagg
780

cccaaggtca aagccaagac caagagccag tga
813

<210> 8

<211> 270

<212> PRT

<213> Homo sapiens

<400> 8

Met Val Thr Ala Met Asn Val Ser His Glu Val Asn Gln Leu Phe Gln
1 5 10 15

Pro Tyr Asn Phe Glu Leu Ser Lys Asp Met Arg Pro Phe Phe Glu Glu
20 25 30

Tyr Trp Ala Thr Ser Phe Pro Ile Ala Leu Ile Tyr Leu Val Leu Ile
35 40 45

Ala Val Gly Gln Asn Tyr Met Lys Glu Arg Lys Gly Phe Asn Leu Gln
50 55 60

Gly Pro Leu Ile Leu Trp Ser Phe Cys Leu Ala Ile Phe Ser Ile Leu
65 70 75 80

Gly Ala Val Arg Met Trp Gly Ile Met Gly Thr Val Leu Leu Thr Gly
85 90 95

Gly Leu Lys Gln Thr Val Cys Phe Ile Asn Phe Ile Asp Asn Ser Thr
100 105 110

Val Lys Phe Trp Ser Trp Val Phe Leu Leu Ser Lys Val Ile Glu Leu
115 120 125

Gly Asp Thr Ala Phe Ile Ile Leu Arg Lys Arg Pro Leu Ile Phe Ile
130 135 140

His Trp Tyr His His Ser Thr Val Leu Val Tyr Thr Ser Phe Gly Tyr
145 150 155 160

Lys Asn Lys Val Pro Ala Gly Gly Trp Phe Val Thr Met Asn Phe Gly
165 170 175

Val His Ala Ile Met Tyr Thr Tyr Tyr Thr Leu Lys Ala Ala Asn Val
180 185 190

Lys Pro Pro Lys Met Leu Pro Met Leu Ile Thr Ser Leu Gln Ile Leu
195 200 205

Gln Met Phe Val Gly Ala Ile Val Ser Ile Leu Thr Tyr Ile Trp Arg
210 215 220

Gln Asp Gln Gly Cys His Thr Thr Met Glu His Leu Phe Trp Ser Phe
225 230 235 240

Ile Leu Tyr Met Thr Tyr Phe Ile Leu Phe Ala His Phe Phe Cys Gln
245 250 255

Thr Tyr Ile Arg Pro Lys Val Lys Ala Lys Thr Lys Ser Gln
260 265 270

<210> 9

<211> 3011

<212> DNA

<213> Homo sapiens

<400> 9
ggcacgagcc aaccccggcc taggctctcc accgcatcgg attctggaat ttacgatcac
60

gaaagttcta ttgtcccgcg attggctccc gggccgcatg acatcatagc gcttgattca
120

tccttcgggt cccgattggc tggccgcgcc attgtgacgt cacggtcagc ccacgttctg
180

attgtagata gccggcgcct tcctcttccc atcgcgcggg tcctagccac cggtgtctcc
240

ttctacatcc gcctctgcgc cggctgccac ccgcgctccc tccgccgccg ccgccttgct
300

gctgctcaaa gctgctgccg ccccttgggc taaaaggttt tcaaatggaa cattttgatg
360

catcacttag tacctatttc aaggcattgc taggccctcg agatactaga gtaaaaggat
420

ggtttcttct ggacaattat atacccacat ttatctgctc tgtcatatat ttactaattg
480

tatggctggg accaaaatac atgaggaata aacagccatt ctcttgccgg gggattttag
540

tggtgtataa ccttggactc acactgctgt ctctgtatat gttctgtgag ttagtaacag
600

gagtatggga aggcaaatac aacttcttct gtcagggcac acgcaccgca ggagaatcag
660

atatgaagat tatccgtgtc ctctggtggt actacttctc caaactcata gaatttatgg
720

acactttctt cttcatcctg cgcaagaaca accaccagat cacggtcctg cacgtctacc
780

accatgcctc gatgctgaac atctggtggt ttgtgatgaa ctgggtcccc tgcggccact
840

cttattttgg tgccacactt aatagcttca tccacgtcct catgtactct tactatggtt
900

tgtcgtcagt cccttccatg cgtccatacc tctggtggaa gaagtacatc actcaggggc
960

agctgcttca gtttgtgctg acaatcatcc agaccagctg cggggtcatc tggccgtgca
1020

cattccctct tggttggttg tatttccaga ttggatacat gatttccctg attgctctct
1080

tcacaaactt ctacattcag acctacaaca agaaaggggc ctcccgaagg aaagaccacc
1140

tgaaggacca ccagaatggg tccatggctg ctgtgaatgg acacaccaac agcttttcac
1200

ccctggaaaa caatgtgaag ccaaggaagc tgcggaagga ttgaagtcaa agaattgaaa
1260

ccctccaaac cacgtcatct gattgtaagc acaatatgag ttgtgcccca atgctcgtta
1320

acagctgctg taactagtct ggcctacaat agtgtgattc atgtaggact tctttcatca
1380

attcaaaacc cctagaaaac gtatacagat tatataagta gggataagat ttctaacatt
1440

tctgggctct ctgacccctg cgctagactg tggaaaggga gtattattat agtatacaac
1500

actgctgttg ccttattagt tataacatga taggtgctga attgtgattc acaatttaaa
1560

aacactgtaa tccaaacttt ttttttaac tgtagatcat gcatgtgatt gtaaatgtaa
1620

atttgtacaa tgttgttatg gtagagaaac acacatgcct taaaatttaa aaagcagggc
1680

ccaaagctta ttagtttaaa ttagggtatg tttcaagttt gtattaattt gtaatagctc
1740

tgtttagaaa aaatcaaaga ccatgattta tgaaactaat gtgacataat ttccagtgac
1800

ttgttgatgt gaaatcagac acggcacctt cagttttgta ctattggctt tgaatcaagc
1860

aggctcaaat ctagtggaac agtcagttta acttttaac agatcttatt tttttatttt
1920

gagtgccact attaatgtaa aaaggggggg gctctacagc agtcgtgatg aaacttaaat
1980

atatattctt tgtcctcgag attttaggaa gggtgtaggg tgagtaggcc atttttaatt
2040

tctgaagtgc taagtgtttt tatacagcaa acaaaaagtc aattttgctt tccaccagtg
2100

cgagagagga tgtatacttt tcaagagaga tgattgccta tttaccgttt gacagagtcc
2160

cgtagatgag caatggggaa ctggttgcca gggtctaaat ttggattgat ttatgcactg
2220

ttatctgttt tgacacagat ttccttgtaa aatgtgccta gtttaccaaa attaacaaag
2280

ggggggaaag gaccttagaa ctttttaagg taaaatcaaa tatagctaca gcataagaga
2340

atcgagaaat ttgatagagg taacttgttt aatgtaaatc taatagtact tgtaatttct
2400

ttctgcttag aatctaaaga tgtgtttaga acctcttgtt taaaaataat agactgctta
2460

tcataaaatc acatctcaca catttgaggc agtggtcaaa caggtaaagc ctatgatgtg
2520

tgtcatttta aagtgtcgga atttagcctc tgaatacctt ctccattggg ggaaagatat
2580

tcttggaacc actcatgaca tatcttagaa ggtcattgac aatgtataaa ctaattgttg
2640

gtttgatatt tatgtaaata tcagtttacc atgctttaat tttgcacatt cgtactatag
2700

ggagcctatt ggttctctat tagtcttgtg ggttttctgt ttgaaaagga gtcatggcat
2760

ctgtttacat ttaccttatc aaacctagaa tgtgtatatt tataaatgta tgtcttcatt
2820

gctaggtact aatttgcaga tgtctttaca tatttcaata cagaaactat aacattcaat
2880

agtgtgctgt caaagtgtgc ttagctcacc tggatatacc tacattgtta aatgtctaaa
2940

cagtaatcat taaaacattt ttgattacct gtgaaaaaaa aaaaaaaaaa aaaaaaaaaa
3000

aaaaaaaaaa a
3011

<210> 10

<211> 900

<212> DNA

<213> Homo sapiens

<400> 10
atggaacatt ttgatgcatc acttagtacc tatttcaagg cattgctagg ccctcgagat
60

actagagtaa aaggatggtt tcttctggac aattatatac ccacatttat ctgctctgtc
120

atatatttac taattgtatg gctgggacca aaatacatga ggaataaaca gccattctct
180

tgccgggggga ttttagtggt gtataacctt ggactcacac tgctgtctct gtatatgttc
240

tgtgagttag taacaggagt atgggaaggc aaatacaact tcttctgtca gggcacacgc
300

accgcaggag aatcagatat gaagattatc cgtgtcctct ggtggtacta cttctccaaa
360

ctcatagaat ttatggacac tttcttcttc atcctgcgca agaacaacca ccagatcacg
420

gtcctgcacg tctaccacca tgcctcgatg ctgaacatct ggtggtttgt gatgaactgg
480

gtcccctgcg gccactctta ttttggtgcc acacttaata gcttcatcca cgtcctcatg
540

tactcttact atggtttgtc gtcagtccct tccatgcgtc catacctctg gtggaagaag
600

tacatcactc aggggcagct gcttcagttt gtgctgacaa tcatccagac cagctgcggg
660

gtcatctggc cgtgcacatt ccctcttggt tggttgtatt tccagattgg atacatgatt
720

tccctgattg ctctcttcac aaacttctac attcagacct acaacaagaa aggggcctcc
780

cgaaggaaag accacctgaa ggaccaccag aatgggtcca tggctgctgt gaatggacac
840

accaacagct tttcacccct ggaaaacaat gtgaagccaa ggaagctgcg gaaggattga
900

<210> 11

<211> 299

<212> PRT

<213> Homo sapiens

<400> 11

Met Glu His Phe Asp Ala Ser Leu Ser Thr Tyr Phe Lys Ala Leu Leu
1 5 10 15

Gly Pro Arg Asp Thr Arg Val Lys Gly Trp Phe Leu Leu Asp Asn Tyr
20 25 30

Ile Pro Thr Phe Ile Cys Ser Val Ile Tyr Leu Leu Ile Val Trp Leu
35 40 45

Gly Pro Lys Tyr Met Arg Asn Lys Gln Pro Phe Ser Cys Arg Gly Ile
50 55 60

```
Leu Val Val Tyr Asn Leu Gly Leu Thr Leu Leu Ser Leu Tyr Met Phe
65                  70                  75                  80

Cys Glu Leu Val Thr Gly Val Trp Glu Gly Lys Tyr Asn Phe Phe Cys
                85                  90                  95

Gln Gly Thr Arg Thr Ala Gly Glu Ser Asp Met Lys Ile Ile Arg Val
            100                 105                 110

Leu Trp Trp Tyr Tyr Phe Ser Lys Leu Ile Glu Phe Met Asp Thr Phe
        115                 120                 125

Phe Phe Ile Leu Arg Lys Asn Asn His Gln Ile Thr Val Leu His Val
    130                 135                 140

Tyr His His Ala Ser Met Leu Asn Ile Trp Trp Phe Val Met Asn Trp
145                 150                 155                 160

Val Pro Cys Gly His Ser Tyr Phe Gly Ala Thr Leu Asn Ser Phe Ile
                165                 170                 175

His Val Leu Met Tyr Ser Tyr Tyr Gly Leu Ser Ser Val Pro Ser Met
            180                 185                 190

Arg Pro Tyr Leu Trp Trp Lys Lys Tyr Ile Thr Gln Gly Gln Leu Leu
        195                 200                 205

Gln Phe Val Leu Thr Ile Ile Gln Thr Ser Cys Gly Val Ile Trp Pro
    210                 215                 220

Cys Thr Phe Pro Leu Gly Trp Leu Tyr Phe Gln Ile Gly Tyr Met Ile
225                 230                 235                 240

Ser Leu Ile Ala Leu Phe Thr Asn Phe Tyr Ile Gln Thr Tyr Asn Lys
                245                 250                 255

Lys Gly Ala Ser Arg Arg Lys Asp His Leu Lys Asp His Gln Asn Gly
            260                 265                 270

Ser Met Ala Ala Val Asn Gly His Thr Asn Ser Phe Ser Pro Leu Glu
        275                 280                 285
```

Asn Asn Val Lys Pro Arg Lys Leu Arg Lys Asp
290 295

<210> 12

<211> 2992

<212> DNA

<213> Homo sapiens

<400> 12

ctcgtctttc tcccgggaac cttgacgacg ccttccgctt ggccctgcct tctgccgcat 60

ccccgccgcc gcggcgcctt gaggagcagg agaagacgca gccgggccgc cgccgttaga 120

ggggttcccg gccgccgctc gccccgtcgg ccgccaccgc ctccggggtc agccctctct 180

ctgggtctcc gctttctcct gccgccagcg cccgctcatc gccgcgatgg ggctcctgga 240

ctcggagccg ggtagtgtcc taaacgtagt gtccacggca ctcaacgaca cggtagagtt 300

ctaccgctgg acctggtcca tcgcagataa gcgtgtggaa aattggcctc tgatgcagtc 360

tccttggcct acactaagta taagcactct ttatctcctg tttgtgtggc tgggtccaaa 420

atggatgaag gaccgagaac cttttcagat gcgtctagtg ctcattatct ataattttgg 480

gatggttttg cttaacctct ttatcttcag agagttattc atgggatcat ataatgcggg 540

atatagctat atttgccaga gtgtggatta ttctaataat gttcatgaag tcaggatagc 600

tgctgctctg tggtggtact ttgtatctaa aggagttgag tatttggaca cagtgttttt 660

tattctgaga aagaaaaaca accaagtttc tttccttcat gtgtatcatc actgtacgat 720

gtttaccttg tggtggattg gaattaagtg ggttgcagga ggacaagcat tttttggagc 780

ccagttgaat tcctttatcc atgtgattat gtactcatac tatgggttaa ctgcatttgg 840

cccatggatt cagaaatatc tttggtggaa acgatacctg actatgttgc aactgattca
900

attccatgtg accattgggc acacggcact gtctctttac actgactgcc ccttccccaa
960

atggatgcac tgggctctaa ttgcctatgc aatcagcttc atatttctct ttcttaactt
1020

ctacattcgg acatacaaag agcctaagaa accaaaagct ggaaaaacag ccatgaatgg
1080

tatttcagca aatggtgtga gcaaatcaga aaaacaactc atgatagaaa atggaaaaaa
1140

gcagaaaaat ggaaaagcaa aaggagatta aattgaactg ggccttaact gttgttgaca
1200

gtgaggaaaa actcccatat catataaaat ttcagggaaa acagaagcaa aggagagctt
1260

gggggtgggg agaaaagaca aatgtgctct atgtcctagt aactcttaga ctgagtaaag
1320

tgttaatacc atacccagat gttttattta tgaagttttt attttaaaca tttttttaa
1380

aaattagcct tgatattctc cagaccaaag caatcattaa gtgactttgg ggattctccc
1440

cctgttcaca tccagttgtc taaaggatga gatttttcat gtatcttata gtcactcatt
1500

cttcggtctg aattttagac gatcacagaa acggtcttta tgaattattt tgataaatta
1560

ctaattatct tatctactga ctgaaatcag tggtgttaca ttttcttgtc caaagctgaa
1620

aatgtgtata cacttaaact tgcacatttg aattcatttg ctgaccggaa tggtcaaatc
1680

tctccacctc tagtcagagt ataattttgg ttgtaattaa atttttaaaa tctgctgatc
1740

tctgtagaat cttagaggct tgatgatgat ggtgttggtg aaaataagaa agaattgcag
1800

taaagtcttg tctggtgacc cagagatcac catgacttga ggcacaaatc actgtgggga
1860

aacaattttt tgtgatgaaa aggcagcatt tgaatactcc tgttagtagc agaaatatat
1920

tatgaaaatt aagattattg tctgattgaa acatgaaaca actcatgtct ttattagtaa
1980

catcataaga tagttacatt tatgtgctgt tagaatatgt tgatttttat caggctttcc
2040

ttgttttgat ttatggctgt tcctgatttt tcatatgtgg aaatatacct acctcttccg
2100

ttggaaagaa catttaaaat taaataaatt ttaattaaaa aatcaaggag tcttctaatg
2160

taaattttaa tgttaacttt caaatccact agtatttttt ttgcttttat gacaaatagc
2220

atacaccaaa catttctgtg aaactatcct tctctttcaa tgtgtttaat tttggagtaa
2280

cgttttcctt gtgactaagt tgcaagatct tatttattaa ctaggtatga agtataaacc
2340

cattttggtg caatattctt gactccttgg tgctaaagat tgttaaattc aatgcttgat
2400

gttacaaggt gttgttaaaa cacaaaatga ataaaagtga gagtagtcag aactataaca
2460

ttcaatttgc tatttacaaa tgaagtattt catgtaatat aagtgaacaa ctggaaataa
2520

agtaggaaag aatttgtatc atgttttact acataggtta atttttaag ggatgttgca
2580

aagggattac tagagaaaga caaaatgtga ccaaaaaaaa gcatgaatat ttcttaagta
2640

tctcaacaac atgtcaaagc tgcatgtgta ggatgtatgc tgtttgtaca aactatttca
2700

gaatattttg taagctataa catatttatt gtgcattaaa attaaatact ttttccccaa
2760

aggcatgcag tcatgagaat tacagaaaat ttgcaacata taaagtagtt tgatctaaga
2820

ggattcaaca cctttgtttt gttgctcagt gtgtaatgac tgagatttgt aaatctttgt
2880

gaacattctg tactggttcc caagagctat tcattccctg ctacctgatt tcagcacaat
2940

aaatatactt ctgctgtggg aaaaataaaa aaaaaaaaaa aaaaaaaaaa aa
2992

<210> 13

<211> 945

<212> DNA

<213> Homo sapiens

<400> 13
atggggctcc tggactcgga gccgggtagt gtcctaaacg tagtgtccac ggcactcaac
60

gacacggtag agttctaccg ctggacctgg tccatcgcag ataagcgtgt ggaaaattgg
120

cctctgatgc agtctccttg gcctacacta agtataagca ctctttatct cctgtttgtg
180

tggctgggtc caaaatggat gaaggaccga gaacctttтc agatgcgtct agtgctcatt
240

atctataatt ttgggatggt tttgcttaac ctctttatct tcagagagtt attcatggga
300

tcatataatg cgggatatag ctatatttgc cagagtgtgg attattctaa taatgttcat
360

gaagtcagga tagctgctgc tctgtggtgg tactttgtat ctaaaggagt tgagtatttg
420

gacacagtgt tttttattct gagaaagaaa aacaaccaag tttctttcct tcatgtgtat
480

catcactgta cgatgtttac cttgtggtgg attggaatta agtgggttgc aggaggacaa
540

gcatttttтg gagcccagtt gaattccttt atccatgtga ttatgtactc atactatggg
600

ttaactgcat ttggcccatg gattcagaaa tatctttggt ggaaacgata cctgactatg
660

ttgcaactga ttcaattcca tgtgaccatt gggcacacgg cactgtctct ttacactgac
720

tgccccttcc ccaaatggat gcactgggct ctaattgcct atgcaatcag cttcatattt
780

ctctttctta acttctacat tcggacatac aaagagccta agaaaccaaa agctggaaaa
840

acagccatga atggtatttc agcaaatggt gtgagcaaat cagaaaaaca actcatgata
900

gaaaatggaa aaaagcagaa aaatggaaaa gcaaaaggag attaa
945

<210> 14

<211> 314

<212> PRT

<213> Homo sapiens

<400> 14

Met Gly Leu Leu Asp Ser Glu Pro Gly Ser Val Leu Asn Val Val Ser
1 5 10 15

Thr Ala Leu Asn Asp Thr Val Glu Phe Tyr Arg Trp Thr Trp Ser Ile
20 25 30

Ala Asp Lys Arg Val Glu Asn Trp Pro Leu Met Gln Ser Pro Trp Pro
35 40 45

Thr Leu Ser Ile Ser Thr Leu Tyr Leu Leu Phe Val Trp Leu Gly Pro
50 55 60

Lys Trp Met Lys Asp Arg Glu Pro Phe Gln Met Arg Leu Val Leu Ile
65 70 75 80

Ile Tyr Asn Phe Gly Met Val Leu Leu Asn Leu Phe Ile Phe Arg Glu
85 90 95

Leu Phe Met Gly Ser Tyr Asn Ala Gly Tyr Ser Tyr Ile Cys Gln Ser
100 105 110

Val Asp Tyr Ser Asn Asn Val His Glu Val Arg Ile Ala Ala Ala Leu
115 120 125

Trp Trp Tyr Phe Val Ser Lys Gly Val Glu Tyr Leu Asp Thr Val Phe
130 135 140

Phe Ile Leu Arg Lys Lys Asn Asn Gln Val Ser Phe Leu His Val Tyr
145 150 155 160

His His Cys Thr Met Phe Thr Leu Trp Trp Ile Gly Ile Lys Trp Val
165 170 175

Ala Gly Gly Gln Ala Phe Phe Gly Ala Gln Leu Asn Ser Phe Ile His

180 185 190

Val Ile Met Tyr Ser Tyr Tyr Gly Leu Thr Ala Phe Gly Pro Trp Ile
195 200 205

Gln Lys Tyr Leu Trp Trp Lys Arg Tyr Leu Thr Met Leu Gln Leu Ile
210 215 220

Gln Phe His Val Thr Ile Gly His Thr Ala Leu Ser Leu Tyr Thr Asp
225 230 235 240

Cys Pro Phe Pro Lys Trp Met His Trp Ala Leu Ile Ala Tyr Ala Ile
245 250 255

Ser Phe Ile Phe Leu Phe Leu Asn Phe Tyr Ile Arg Thr Tyr Lys Glu
260 265 270

Pro Lys Lys Pro Lys Ala Gly Lys Thr Ala Met Asn Gly Ile Ser Ala
275 280 285

Asn Gly Val Ser Lys Ser Glu Lys Gln Leu Met Ile Glu Asn Gly Lys
290 295 300

Lys Gln Lys Asn Gly Lys Ala Lys Gly Asp
305 310

<210> 15

<211> 2340

<212> DNA

<213> Homo sapiens

<400> 15
gatagcgccg ggcagaggga cccggctacc ctggacagcg catcgccgcc cgcccgggtc
60

gccgcgccac agccgctgcg gatcatggaa catctaaagg cctttgatga tgaaatcaat
120

gctttttgg acaatatgtt tggaccgcga gattctcgag tcagagggtg gttcacgttg
180

gactcttacc ttcctacctt ttttcttact gtcatgtatc tgctctcaat atggctgggt
240

aacaagtata tgaagaacag acctgctctt tctctcaggg gtatcctcac cttgtataat
300

cttggaatca cacttctctc cgcgtacatg ctggcagagc tcattctctc cacttgggaa
360

ggaggctaca acttacagtg tcaagatctt accagcgcag gggaagctga catccgggta
420

gccaaggtgc tttggtggta ctatttctcc aaatcagtag agttcctgga cacaattttc
480

ttcgttttgc ggaaaaaaac gagtcagatt acttttcttc atgtatatca tcatgcttct
540

atgtttaaca tctggtggtg tgtcttgaac tggatacctt gtggacaaag tttctttgga
600

ccaacactga acagttttgt ccacattctt atgtactcct actatggact ttctgtgttt
660

ccatctatgc acaagtatct ttggtggaag aaatatctca cacaggctca gctggtgcag
720

ttcgtgctca ccatcacgca caccatgagc gccgtcgtga aaccgtgtgg cttccccttc
780

ggttgtctca tcttccagtc atcttatatg ctaacgttag tcatcctctt cttaaatttt
840

tatgttcaga cataccgaaa aaagccaatg aagaaagata tgcaagagcc acctgcaggg
900

aaagaagtga agaatggttt ttccaaagcc tacttcactg cagcaaatgg agtgatgaac
960

aagaaagcac aataaaaatg agtaacagaa aaagcacata tactagccta acagattggc
1020

ttgttttaaa gcaaagactg aattgaaggt tacatgtttt aggataaact aatttctttt
1080

gagttcataa atcatttgta cccagaatgt attaatatat tgctattagg ttaatctgtt
1140

aactgaatgc tttgatcagc attgaggtga tgctcacctc cgaggacctc agaactggtg
1200

cagcttctct ctccctccct cccacagact gaacctttcg ccagaagctg tccttataac
1260

gccttatacg catacacagc caggaaacgt ggagcattgt ttctcacaga gagtctccaa
1320

ataaaaaggg ttttgttcag attaaaatgt ttacaacaaa atgttaatta tattctaaat
1380

acagggtatg ttctaatcta tattaagcaa taatgccagt gcataatcat tccatttgtt
1440

cctttagcaa tcaaccccag aaaatattaa aatgggatca tacacagaag atagaaaaat
1500

ctagcaaaac ttctctttct gtaagccaga gtcttgtcta tcagattccc acaaccactc
1560

ctgattctaa atttagtgat atggtaatga aattggtatt tattttaaat attagttatt
1620

ctaaggagaa aaaaatgctt ctgcaagatt ttcataattc aggggctgtg gataggattg
1680

ttcctctgtt tccctaatca ttcatctgtt catgtctccc tcttgtgcca gtcagcctag
1740

gttatacaga tgccatgctc cacaccacga gcagtgtaca aatctggctg cccgtttact
1800

ttctgagcaa gcactggagt ccactccgac ctttttcttt gaacatgcat gctgctggaa
1860

tatgtataaa tcagaactag cagaagtagc agagtgatgg gagcaaaata ggcactgaat
1920

tcgtcaactc ttttttgtga gcctacttgt gaatattacc tcagatacct gttgtcactc
1980

ttcacaggtt atttaagttc ttgaagctgg gaggaaaaag atggagtagc ttggaaagat
2040

tccagcactg agccgtgagc cggtcatgag ccacgataaa aaatgccagt ttggcaaact
2100

cagcactcct gttccctgct caggtatatg cgatctctac tgagaagcaa gcacaaaagt
2160

agaccaaagt attaatgagt atttcctttc tccataagtg caggactgtt actcactact
2220

aaactctacc aagaatggaa accaagaata ttttctgaag attttttttga agattaattt
2280

ataccctata aaataaaact tgttagcttc gatgaagtca aaaaaaaaaa aaaaaaaaaa
2340

<210> 16

<211> 891

<212> DNA

<213> Homo sapiens

<400> 16
atggaacatc taaaggcctt tgatgatgaa atcaatgctt ttttggacaa tatgtttgga
60

ccgcgagatt ctcgagtcag agggtggttc acgttggact cttaccttcc tacctttttt
120

cttactgtca tgtatctgct ctcaatatgg ctgggtaaca agtatatgaa gaacagacct
180

gctctttctc tcaggggtat cctcaccttg tataatcttg gaatcacact tctctccgcg
240

tacatgctgg cagagctcat tctctccact tgggaaggag gctacaactt acagtgtcaa
300

gatcttacca gcgcagggga agctgacatc cgggtagcca aggtgctttg gtggtactat
360

ttctccaaat cagtagagtt cctggacaca attttcttcg ttttgcggaa aaaaacgagt
420

cagattactt ttcttcatgt atatcatcat gcttctatgt ttaacatctg gtggtgtgtc
480

ttgaactgga taccttgtgg acaaagtttc tttggaccaa cactgaacag ttttgtccac
540

attcttatgt actcctacta tggactttct gtgtttccat ctatgcacaa gtatctttgg
600

tggaagaaat atctcacaca ggctcagctg gtgcagttcg tgctcaccat cacgcacacc
660

atgagcgccg tcgtgaaacc gtgtggcttc cccttcggtt gtctcatctt ccagtcatct
720

tatatgctaa cgttagtcat cctcttctta aatttttatg ttcagacata ccgaaaaaag
780

ccaatgaaga aagatatgca agagccacct gcagggaaag aagtgaagaa tggtttttcc
840

aaagcctact tcactgcagc aaatggagtg atgaacaaga aagcacaata a
891

<210> 17

<211> 296

<212> PRT

<213> Homo sapiens

<400> 17

```
Met Glu His Leu Lys Ala Phe Asp Asp Glu Ile Asn Ala Phe Leu Asp
1               5                  10                  15

Asn Met Phe Gly Pro Arg Asp Ser Arg Val Arg Gly Trp Phe Thr Leu
            20                  25                  30

Asp Ser Tyr Leu Pro Thr Phe Phe Leu Thr Val Met Tyr Leu Leu Ser
        35                  40                  45

Ile Trp Leu Gly Asn Lys Tyr Met Lys Asn Arg Pro Ala Leu Ser Leu
    50                  55                  60

Arg Gly Ile Leu Thr Leu Tyr Asn Leu Gly Ile Thr Leu Leu Ser Ala
65                  70                  75                  80

Tyr Met Leu Ala Glu Leu Ile Leu Ser Thr Trp Glu Gly Gly Tyr Asn
                85                  90                  95

Leu Gln Cys Gln Asp Leu Thr Ser Ala Gly Glu Ala Asp Ile Arg Val
            100                 105                 110

Ala Lys Val Leu Trp Trp Tyr Tyr Phe Ser Lys Ser Val Glu Phe Leu
        115                 120                 125

Asp Thr Ile Phe Phe Val Leu Arg Lys Lys Thr Ser Gln Ile Thr Phe
    130                 135                 140

Leu His Val Tyr His His Ala Ser Met Phe Asn Ile Trp Trp Cys Val
145                 150                 155                 160

Leu Asn Trp Ile Pro Cys Gly Gln Ser Phe Phe Gly Pro Thr Leu Asn
                165                 170                 175

Ser Phe Val His Ile Leu Met Tyr Ser Tyr Tyr Gly Leu Ser Val Phe
            180                 185                 190

Pro Ser Met His Lys Tyr Leu Trp Trp Lys Lys Tyr Leu Thr Gln Ala
        195                 200                 205
```

Gln Leu Val Gln Phe Val Leu Thr Ile Thr His Thr Met Ser Ala Val
210 215 220

Val Lys Pro Cys Gly Phe Pro Phe Gly Cys Leu Ile Phe Gln Ser Ser
225 230 235 240

Tyr Met Leu Thr Leu Val Ile Leu Phe Leu Asn Phe Tyr Val Gln Thr
245 250 255

Tyr Arg Lys Lys Pro Met Lys Lys Asp Met Gln Glu Pro Pro Ala Gly
260 265 270

Lys Glu Val Lys Asn Gly Phe Ser Lys Ala Tyr Phe Thr Ala Ala Asn
275 280 285

Gly Val Met Asn Lys Lys Ala Gln
290 295

<210> 18

<211> 1561

<212> DNA

<213> Homo sapiens

<400> 18
gcccagcaga tgaggaagtg gcaggcaggc aggctggccc cggggacttc tctctggccc
60

tgctccctcc gagcgctccg ccgttgcccg cctggcccct acggagtcct tagccaggat
120

ggaggctgtt gtgaacttgt accaagaggt gatgaagcac gcagatcccc ggatccaggg
180

ctaccctctg atggggtccc ccttgctaat gacctccatt ctcctgacct acgtgtactt
240

cgttctctca cttgggcctc gcatcatggc taatcggaag cccttccagc tccgtggctt
300

catgattgtc tacaacttct cactggtggc actctccctc tacattgtct atgagttcct
360

gatgtcgggc tggctgagca cctatacctg gcgctgtgac cctgtggact attccaacag
420

ccctgaggca cttaggatgg ttcgggtggc ctggctcttc ctcttctcca agttcattga
480

gctgatggac acagtgatct ttattctccg aaagaaagac gggcaggtga ccttcctaca
540

tgtcttccat cactctgtgc ttccctggag ctggtggtgg ggggtaaaga ttgccccggg
600

aggaatgggc tctttccatg ccatgataaa ctcttccgtg catgtcataa tgtacctgta
660

ctacggatta tctgcctttg gccctgtggc acaaccctac ctttggtgga aaaagcacat
720

gacagccatt cagctgatcc agtttgtcct ggtctcactg cacatctccc agtactactt
780

tatgtccagc tgtaactacc agtacccagt cattattcac ctcatctgga tgtatggcac
840

catcttcttc atgctgttct ccaacttctg gtatcactct tataccaagg gcaagcggct
900

gccccgtgca cttcagcaaa atggagctcc aggtattgcc aaggtcaagg ccaactgaga
960

agcatggcct agataggcgc ccacctaagt gcctcaggac tgcaccttag ggcagtgtcc
1020

gtcagtgccc tctccaccta cacctgtgac caaggcttat gtggtcagga ctgagcaggg
1080

gactggccct cccctcccca cagctgctct acagggacca cggctttggt tcctcaccca
1140

cttcccccgg gcagctccag ggatgtggcc tcattgctgt ctgccactcc agagctgggg
1200

gctaaaaggg ctgtacagtt atttcccccc ccctgcctta aaacttggga gaggagcact
1260

cagggctggc cccacaaagg gtctcgtggc ctttttcctc acacagaaga ggtcagcaat
1320

aatgtcactg tggacccagt ctcactcctc caccccacac actgaagcag tagcttctgg
1380

gccaaaggtc agggtgggcg ggggcctggg aatacagcct gtggaggctg cttactcaac
1440

ttgtgtctta attaaaagtg acagaggaaa ccacggaggc tgtgtgtata cgtatgtgaa
1500

cagaggaggt gggggaacac aacacgctta gccctgtcca caaacaagag ggatggatgg
1560

t
1561

<210> 19

<211> 840

<212> DNA

<213> Homo sapiens

<400> 19
atggaggctg ttgtgaactt gtaccaagag gtgatgaagc acgcagatcc ccggatccag
60

ggctaccctc tgatggggtc ccccttgcta atgacctcca ttctcctgac ctacgtgtac
120

ttcgttctct cacttgggcc tcgcatcatg gctaatcgga agcccttcca gctccgtggc
180

ttcatgattg tctacaactt ctcactggtg gcactctccc tctacattgt ctatgagttc
240

ctgatgtcgg gctggctgag cacctatacc tggcgctgtg accctgtgga ctattccaac
300

agccctgagg cacttaggat ggttcgggtg gcctggctct tcctcttctc caagttcatt
360

gagctgatgg acacagtgat ctttattctc cgaaagaaag acgggcaggt gaccttccta
420

catgtcttcc atcactctgt gcttccctgg agctggtggt ggggggtaaa gattgccccg
480

ggaggaatgg gctctttcca tgccatgata aactcttccg tgcatgtcat aatgtacctg
540

tactacggat tatctgcctt tggccctgtg gcacaaccct acctttggtg gaaaaagcac
600

atgacagcca ttcagctgat ccagtttgtc ctggtctcac tgcacatctc ccagtactac
660

tttatgtcca gctgtaacta ccagtaccca gtcattattc acctcatctg gatgtatggc
720

accatcttct tcatgctgtt ctccaacttc tggtatcact cttataccaa gggcaagcgg
780

ctgccccgtg cacttcagca aaatggagct ccaggtattg ccaaggtcaa ggccaactga
840

<210> 20

<211> 279

<212> PRT

<213> Homo sapiens

<400> 20

Met Glu Ala Val Val Asn Leu Tyr Gln Glu Val Met Lys His Ala Asp
1 5 10 15

Pro Arg Ile Gln Gly Tyr Pro Leu Met Gly Ser Pro Leu Leu Met Thr
20 25 30

Ser Ile Leu Leu Thr Tyr Val Tyr Phe Val Leu Ser Leu Gly Pro Arg
35 40 45

Ile Met Ala Asn Arg Lys Pro Phe Gln Leu Arg Gly Phe Met Ile Val
50 55 60

Tyr Asn Phe Ser Leu Val Ala Leu Ser Leu Tyr Ile Val Tyr Glu Phe
65 70 75 80

Leu Met Ser Gly Trp Leu Ser Thr Tyr Thr Trp Arg Cys Asp Pro Val
85 90 95

Asp Tyr Ser Asn Ser Pro Glu Ala Leu Arg Met Val Arg Val Ala Trp
100 105 110

Leu Phe Leu Phe Ser Lys Phe Ile Glu Leu Met Asp Thr Val Ile Phe
115 120 125

Ile Leu Arg Lys Lys Asp Gly Gln Val Thr Phe Leu His Val Phe His
130 135 140

His Ser Val Leu Pro Trp Ser Trp Trp Trp Gly Val Lys Ile Ala Pro
145 150 155 160

Gly Gly Met Gly Ser Phe His Ala Met Ile Asn Ser Ser Val His Val
165 170 175

Ile Met Tyr Leu Tyr Tyr Gly Leu Ser Ala Phe Gly Pro Val Ala Gln
180 185 190

Pro Tyr Leu Trp Trp Lys Lys His Met Thr Ala Ile Gln Leu Ile Gln
195 200 205

Phe Val Leu Val Ser Leu His Ile Ser Gln Tyr Tyr Phe Met Ser Ser
210 215 220

Cys Asn Tyr Gln Tyr Pro Val Ile Ile His Leu Ile Trp Met Tyr Gly
225 230 235 240

Thr Ile Phe Phe Met Leu Phe Ser Asn Phe Trp Tyr His Ser Tyr Thr
245 250 255

Lys Gly Lys Arg Leu Pro Arg Ala Leu Gln Gln Asn Gly Ala Pro Gly
260 265 270

Ile Ala Lys Val Lys Ala Asn
275

<210> 21

<211> 2393

<212> DNA

<213> Homo sapiens

<400> 21
ggccggcgcc tcctcctgga ttcattcact cgctcttttc attcacgaag gtagtgaggc
60

ctagtggaaa gccatggaga gcgctctccc cgccgccggc ttcctgtact gggtcggcgc
120

gggcaccgtg gcctacctag ccctgcgtat ttcgtactcg ctcttcacgg ccctccgggt
180

ctggggagtg gggaatgagg cgggggtcgg cccggggctc ggagagtggg cagttgtcac
240

aggtagtact gatggaattg gaaaatcata tgcagaagag ttagcaaagc atggaatgaa
300

ggttgtcctt atcagcagat caaaggataa acttgaccag gtttccagtg aaataaaaga
360

aaaattcaaa gtggagacaa gaaccattgc tgttgacttt gcatcagaag atatttatga
420

taaaattaaa acaggcttgg ctggtcttga aatcggcatc ttagtgaaca acgtgggaat
480

gtcgtatgag tatcctgaat actttttgga tgttcctgac ttggacaatg tgatcaagaa
540

aatgataaat attaatattc tttctgtttg taagatgaca caattggtac tgcctggcat
600

ggtggaaaga tccaaagggg ctattctgaa catttcatct ggcagtggca tgctccctgt
660

cccactcttg accatctatt ctgcaaccaa gacttttgta gatttcttct ctcagtgcct
720

ccatgaggag tataggagca agggcgtctt tgtgcagagt gtcctgccat acttcgtagc
780

tacaaaactg gctaaaatcc ggaagccaac tttggataag ccctctccgg agacgtttgt
840

gaagtctgca attaaaacag tcggcctgca atcccgaacc aatggatacc tgatccatgc
900

tcttatgggc tcgataatct caaacctgcc ttcttggatt tatttgaaaa tagtcatgaa
960

tatgaacaag tctacacggg ctcactatct gaagaaaacc aagaagaact aagcatgata
1020

actgcattgt aacttggcca gatgctccag catatgcacg ttcactgcaa agcaccctac
1080

tggttttgaa aatctgacct tgtcatttca atagttatta acatgactaa atattatctt
1140

aattaagagg aaaatagaag ttgcttttag gggtttctga catatattct ggatactatc
1200

cgaggtaatt ttgaagttta atataaatgc tcatatcaaa tgaatataga actaatattg
1260

tcgggaacac ctaatagaaa ggaatactat tatagcaaat cacagaatga tagactcaag
1320

cataaaactt ggcagtttta tctgcttcaa aatgccattg atcattattc ctgtattttc
1380

tctgaaactg attataaaaa ccaatgtcca gctactcttt tgtttttgac acttgaagaa
1440

atggagatcg atttgatttg tttataagca gacacactgc aatttacaaa gatctcttta
1500

cggttttata aaattatctt ccagtttgta catttatatg gaattgttct ttatcaaggg
1560

tagctaatga catgaaaata attgtgaaat atggaattat ttctgacaca tgaagcccac
1620

taaactatgc tttcctataa tgcatatttc ttctcagttt aaatgtatgt aaatatcgaa
1680

gctatatggt atgatttata aagataaatg ggccaaagtg tacattgaga ctggcagcca
1740

tctatggtac cactgaaacc ctgacccaga aaagtggctt gcttggacac ccagctgcct
1800

ttgtttctgc attaaaccaa tattgatcac acatatgaca caggctagtc ctataaaagt
1860

aatgacttca tagaaatggc attataattt ttaagttgat actctacagg tagctattga
1920

tataattagt tttaataaaa catgctgcaa ccatggtata caacaaaaat acatttcttt
1980

ggtgattgaa attaaggccg tatttacaat gacttaatat aagactgact tttatcctgc
2040

ttcataactt gtatggagaa ctcaccaaga aagaattcaa tactgtgaaa tatgcagcaa
2100

gaagattggt ctttacctag gctgtgtttc ctaagctctg agttttcagc accagtagat
2160

ttgtattaaa agaaaaaaaa atggggcctt agcttctggc ttttaatttt gccagctaag
2220

gacataaaac aaaaataaac aaacaaaaac aaatagccat ctgctatcag catcattatg
2280

taaaagaaaa tatattttag cccctaaaat taggaagaat gtaatctcag aataaaggtt
2340

gtcatttaag ttgaataaat atatagcttt atgaaaaaca caaaaaaaaa aaa
2393

<210> 22

<211> 939

<212> DNA

<213> Homo sapiens

<400> 22

atggagagcg ctctccccgc cgccggcttc ctgtactggg tcggcgcggg caccgtggcc
60

tacctagccc tgcgtatttc gtactcgctc ttcacggccc tccgggtctg gggagtgggg
120

aatgaggcgg gggtcggccc ggggctcgga gagtgggcag ttgtcacagg tagtactgat
180

ggaattggaa aatcatatgc agaagagtta gcaaagcatg gaatgaaggt tgtccttatc
240

agcagatcaa aggataaact tgaccaggtt tccagtgaaa taaaagaaaa attcaaagtg
300

gagacaagaa ccattgctgt tgactttgca tcagaagata tttatgataa aattaaaaca
360

ggcttggctg gtcttgaaat cggcatctta gtgaacaacg tgggaatgtc gtatgagtat
420

cctgaatact ttttggatgt tcctgacttg gacaatgtga tcaagaaaat gataaatatt
480

aatattcttt ctgtttgtaa gatgacacaa ttggtactgc ctggcatggt ggaaagatcc
540

aaaggggcta ttctgaacat ttcatctggc agtggcatgc tccctgtccc actcttgacc
600

atctattctg caaccaagac ttttgtagat ttcttctctc agtgcctcca tgaggagtat
660

aggagcaagg gcgtctttgt gcagagtgtc ctgccatact tcgtagctac aaaactggct
720

aaaatccgga agccaacttt ggataagccc tctccggaga cgtttgtgaa gtctgcaatt
780

aaaacagtcg gcctgcaatc ccgaaccaat ggatacctga tccatgctct tatgggctcg
840

ataatctcaa acctgccttc ttggatttat ttgaaaatag tcatgaatat gaacaagtct
900

acacgggctc actatctgaa gaaaaccaag aagaactaa
939

<210> 23

<211> 312

<212> PRT

<213> Homo sapiens

<400> 23

Met Glu Ser Ala Leu Pro Ala Ala Gly Phe Leu Tyr Trp Val Gly Ala
1 5 10 15

Gly Thr Val Ala Tyr Leu Ala Leu Arg Ile Ser Tyr Ser Leu Phe Thr
20 25 30

Ala Leu Arg Val Trp Gly Val Gly Asn Glu Ala Gly Val Gly Pro Gly
35 40 45

Leu Gly Glu Trp Ala Val Val Thr Gly Ser Thr Asp Gly Ile Gly Lys
50 55 60

Ser Tyr Ala Glu Glu Leu Ala Lys His Gly Met Lys Val Val Leu Ile
65 70 75 80

Ser Arg Ser Lys Asp Lys Leu Asp Gln Val Ser Ser Glu Ile Lys Glu
85 90 95

Lys Phe Lys Val Glu Thr Arg Thr Ile Ala Val Asp Phe Ala Ser Glu
100 105 110

Asp Ile Tyr Asp Lys Ile Lys Thr Gly Leu Ala Gly Leu Glu Ile Gly
115 120 125

Ile Leu Val Asn Asn Val Gly Met Ser Tyr Glu Tyr Pro Glu Tyr Phe
130 135 140

Leu Asp Val Pro Asp Leu Asp Asn Val Ile Lys Lys Met Ile Asn Ile
145 150 155 160

Asn Ile Leu Ser Val Cys Lys Met Thr Gln Leu Val Leu Pro Gly Met
165 170 175

Val Glu Arg Ser Lys Gly Ala Ile Leu Asn Ile Ser Ser Gly Ser Gly
180 185 190

Met Leu Pro Val Pro Leu Leu Thr Ile Tyr Ser Ala Thr Lys Thr Phe
195 200 205

Val Asp Phe Phe Ser Gln Cys Leu His Glu Glu Tyr Arg Ser Lys Gly

210 215 220

Val Phe Val Gln Ser Val Leu Pro Tyr Phe Val Ala Thr Lys Leu Ala
225 230 235 240

Lys Ile Arg Lys Pro Thr Leu Asp Lys Pro Ser Pro Glu Thr Phe Val
245 250 255

Lys Ser Ala Ile Lys Thr Val Gly Leu Gln Ser Arg Thr Asn Gly Tyr
260 265 270

Leu Ile His Ala Leu Met Gly Ser Ile Ile Ser Asn Leu Pro Ser Trp
275 280 285

Ile Tyr Leu Lys Ile Val Met Asn Met Asn Lys Ser Thr Arg Ala His
290 295 300

Tyr Leu Lys Lys Thr Lys Lys Asn
305 310

<210> 24

<211> 629

<212> DNA

<213> Homo sapiens

<400> 24
tgggagggag ccatgaagca ttacgaggtg gagattctgg acgcaaagac aagggagaag
60

ctgtgtttct tcgccgcgtg gatggcctat tacatcaatc accctctcta cactcccccc
120

acctacggag ctcagcaggt gaaactggcg ctcgccatct ttgtgatctg ccagctcggc
180

aacttctcca tccacatggc cctgcgggac ctgcggcccg ctgggtccaa gacgcggaag
240

atcccatacc ccaccaagaa ccccttcacg tggctcttcc tgctggtgtc ctgccccaac
300

tgcacctacg aggtggggtc ctggatcggt ttcgccatca tgacgcagtg tctcccagtg
360

gccctgttct ccctggtggg cttcacccag atgaccatct gggccaaggg caagcaccgc
420

agctacctga aggagttccg ggactacccg cccctgcgca tgcccatcat ccccttcctg
480

ctctgagcag ctcacccctg ctgaggctca gcccctcaac ccggtggcat tctgggggag
540

gagtggggcc cacagctctc cagcacccgg aataaagccc gcctgtcccc agtcggaaca
600

aaaaaaaaaa aaaaaaaaaa aaaaaaaaa
629

<210> 25

<211> 474

<212> DNA

<213> Homo sapiens

<400> 25
atgaagcatt acgaggtgga gattctggac gcaaagacaa gggagaagct gtgtttcttc
60

gccgcgtgga tggcctatta catcaatcac cctctctaca ctccccctac ctacggagct
120

cagcaggtga aactggcgct cgccatcttt gtgatctgcc agctcggcaa cttctccatc
180

cacatggccc tgcgggacct gcggcccgct gggtccaaga cgcggaagat cccatacccc
240

accaagaacc ccttcacgtg gctcttcctg ctggtgtcct gccccaactg cacctacgag
300

gtggggtcct ggatcggttt cgccatcatg acgcagtgtc tcccagtggc cctgttctcc
360

ctggtgggct tcacccagat gaccatctgg gccaagggca agcaccgcag ctacctgaag
420

gagttccggg actacccgcc cctgcgcatg cccatcatcc ccttcctgct ctga
474

<210> 26

<211> 157

<212> PRT

<213> Homo sapiens

<400> 26

Met Lys His Tyr Glu Val Glu Ile Leu Asp Ala Lys Thr Arg Glu Lys
1 5 10 15

Leu Cys Phe Phe Ala Ala Trp Met Ala Tyr Tyr Ile Asn His Pro Leu
20 25 30

Tyr Thr Pro Pro Thr Tyr Gly Ala Gln Gln Val Lys Leu Ala Leu Ala
35 40 45

Ile Phe Val Ile Cys Gln Leu Gly Asn Phe Ser Ile His Met Ala Leu
50 55 60

Arg Asp Leu Arg Pro Ala Gly Ser Lys Thr Arg Lys Ile Pro Tyr Pro
65 70 75 80

Thr Lys Asn Pro Phe Thr Trp Leu Phe Leu Leu Val Ser Cys Pro Asn
85 90 95

Cys Thr Tyr Glu Val Gly Ser Trp Ile Gly Phe Ala Ile Met Thr Gln
100 105 110

Cys Leu Pro Val Ala Leu Phe Ser Leu Val Gly Phe Thr Gln Met Thr
115 120 125

Ile Trp Ala Lys Gly Lys His Arg Ser Tyr Leu Lys Glu Phe Arg Asp
130 135 140

Tyr Pro Pro Leu Arg Met Pro Ile Ile Pro Phe Leu Leu
145 150 155

<210> 27

<211> 1115

<212> DNA

<213> artificial sequence

<220>

<223> RNAi fragment

<400> 27
tgccagtgct tcttggttgg agccggatac gtggctcttg cagctgtggc ttatcgtctt
60

ttgacgattt tctcgaatat tttgggccca tacgttcttc tgtcgccaat cgatttgaag
120

aaaagagctg gagcttcttg ggctggttcg tttctatatt tttactctgg tgatgtctta
180

tcgtatagca ctcagagaac tacttttggt tctctggaac tgttatcatt ttcaattgat
240

aataaaattg caatgaaaaa attgttgcgt aaatataaaa taattgtttt gaatttcaaa
300

caaaaaatac atttttcagt tgtcaccgga gccactgacg gaatcggaaa agcatacgcc
360

ttcgaattgg ctcgtcgtgg attcaatgtc ctgctcgttt cgcgtaccca atcaaaactc
420

gatgagacga agaaggagat tctcgagaag tattccagca ttgaggtccg cactgccgcc
480

ttcgacttca ccaacgctgc tccttctgct tacaaagatc ttctcgccac cttgaaccaa
540

gtagagatcg gagttcttat taacaacgtt ggaatgagct acgaatatcc agatgtactt
600

cacaaagttg acggtggaat cgagcgtctt gcaaacatca ccaccatcaa cactcttcca
660

ccaacattgg tgagttttat tagaattgta ctattgtgat cttactgaaa aagttcttaa
720

ctcagctctc cgccggaatc cttccacaaa tggtcgcacg aaaggctgga gtcattgtta
780

atgttggatc ttcagctgga gcaaatcaaa tggctctctg ggctgtgtat tcagctacaa
840

aggtttgaat ttcaaaatat cttccgtctt ttaattataa ttttattcca gaagtatgtc
900

tcctggctca ccgctatcct ccgaaaagaa tatgaacatc aaggaatcac tgtccaaact
960

attgctccaa tgatggtcgc cacaaagatg tcaaaagtca agagaacttc attcttcact
1020

ccagacggag ccgtgttcgc taaatcagct ctgaacactg ttggaaatac ctcagacacc
1080

accggataca tcacgcatca acttcaactc gagct
1115

<210> 28

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> primer

<400> 28
tgccagtgct tcttggttgg
20

<210> 29

<211> 22

<212> DNA

<213> artificial sequence

<220>

<223> primer

<400> 29
agctcgagtt gaagttgatg cg
22

<210> 30

<211> 951

<212> DNA

<213> Caenorhabditis elegans

<400> 30

atggcttgcc agtgcttctt ggttggagcc ggatacgtgg ctcttgcagc tgtggcttat
60

cgtcttttga cgattttctc gaatattttg ggcccatacg ttcttctgtc gccaatcgat
120

ttgaagaaaa gagctggagc ttcttgggct gttgtcaccg gagccactga cggaatcgga
180

aaagcatacg ccttcgaatt ggctcgtcgt ggattcaatg tcctgctcgt ttcgcgtacc
240

caatcaaaac tcgatgagac gaagaaggag attctcgaga agtattccag cattgaggtc
300

cgcactgccg ccttcgactt caccaacgct gctccttctg cttacaaaga tcttctcgcc
360

accttgaacc aagtagagat cggagttctt attaacaacg ttggaatgag ctacgaatat
420

ccagatgtac ttcacaaagt tgacggtgga atcgagcgtc ttgcaaacat caccaccatc
480

aacactcttc caccaacatt gctctccgcc ggaatccttc cacaaatggt cgcacgaaag
540

gctggagtca ttgttaatgt tggatcttca gctggagcaa atcaaatggc tctctgggct
600

gtgtattcag ctacaaagaa gtatgtctcc tggctcaccg ctatcctccg aaaagaatat
660

gaacatcaag gaatcactgt ccaaactatt gctccaatga tggtcgccac aaagatgtca
720

aaagtcaaga gaacttcatt cttcactcca gacggagccg tgttcgctaa atcagctctg
780

aacactgttg gaaatacctc agacaccacc ggatacatca cgcatcaact tcaactcgag
840

ctcatggatc tcattccaac attcatccgc gacaagatcc tcacaaatat gagtgtcgga
900

actcgtgctg ctgctctcag aaagaaggaa agagaagcca aatctcagta a
951

<210> 31

<211> 3488

<212> DNA

<213> Homo sapiens

<400> 31
agcagctccg gcggccgaga cgggggcggc ggccgcgcgg gtctggcggg accggtttgg
60

aagactttgc cggcctgcag attggcctta agagaaggac ggagccacat actgctgacg
120

gcccagaact ggcagagaga aggttgccat ggctgctgtt gacagtttct acctcttgta
180

cagggaaatc gccaggtctt gcaattgcta tatggaagct ctagctttgg ttggagcctg
240

gtatacggcc agaaaaagca tcactgtcat ctgtgacttt tacagcctga tcaggctgca
300

ttttatcccc cgcctgggga gcagagcaga cttgatcaag cagtatggaa gatgggccgt
360

tgtcagcggt gcaacagatg ggattggaaa agcctacgct gaagagttag caagccgagg
420

tctcaatata atcctgatta gtcggaacga ggagaagttg caggttgttg ctaaagacat
480

agccgacacg tacaaagtgg aaactgatat tatagttgcg gacttcagca gcggtcgtga
540

gatctacctt ccaattcgag aagccctgaa ggacaaagac gttggcatct tggtaaataa
600

cgtgggtgtg ttttatccct acccgcagta tttcactcag ctgtccgagg acaagctctg
660

ggacatcata aatgtgaaca ttgccgccgc tagtttgatg gtccatgttg tgttaccggg
720

aatggtggag agaaagaaag gtgccatcgt cacgatctct tctggctcct gctgcaaacc
780

cactcctcag ctggctgcat tttctgcttc taaggcttat ttagaccact tcagcagagc
840

cttgcaatat gaatatgcct ctaaaggaat ctttgtacag agtctaatcc ctttctatgt
900

agccaccagc atgacagcac ccagcaactt tctgcacagg tgctcgtggt tggtgccttc
960

gccaaaagtc tatgcacatc atgctgtttc tactcttggg atttccaaaa ggaccacagg
1020

atattggtcc cattctattc agtttctttt tgcacagtat atgcctgaat ggctctgggt
1080

gtggggagca aatattctca accgttcact acgtaaggaa gccttatgct gcacagcctg
1140

agtctggatg gccacttgag aagttttgcc aactcctggg aacctcgata ttctgacatt
1200

tggaaaaaca catttaattt atctcctgtg tttcattgct gattattcag catactgttg
1260

attcgtcatt tgcaaaacac acataatacc gtcagagtgc tgtgaaaaac cttaagggtg
1320

tgtggatggc acaggatcaa taatgcctga ggctgattga cgacatctac atttcagtgc
1380

tttttcccta agctgtttga aagttacgct tttctgttgt tctagagcca cagcagtcta
1440

atattgaaat ataatatgat ttgtcaggtc ttataatttc agatgttgtt ttttaaggga
1500

aattgaccat ttcactagag gagttgtgct ggtttttaaa tgtgcatcaa gaaagactac
1560

tgaaaagtat tattttgtaa ctaagattgc tggtactatt aggaaaaatc tgtgtgtatc
1620

gtatagctct agctgtttga ctatctgtaa tgaaaatgct gcacttcaac tggtatttca
1680

ttagagaacc gtgtgtgtgc gtgtgtgtgg tgcctttgag caactttatt tatggttacc
1740

atatttttaa aaagattttt tgtcaggatg acttaacatg gactcttata gggtattaaa
1800

acaatctaga ttattccttt tcatcctaaa taagcctacc aaatttcatg ctgttggttt
1860

gccatgaatg atattacttc ctacattata tttgtgtttt ttcaaatctg ctatggaatg
1920

aacttattcc tagatttgga tatgtaagag aaacctgcag tcatcttttg atttataagg
1980

caattcttgt ggataaatag tgatttctca gcctctgacc cattttataa ctgaaattta
2040

gcccttaga gcttgttata tctggttttc ctacgttttt ctatgtaata ttattccatt
2100

ccagtagcat tattgataga aatagtaagt atttatggaa tagtaaaata tggacaaatt
2160

acgtgtgtga catatctgtc aaaataagtt agaagcttat tcttggtttg tgtaatgaat
2220

ttatgtattg tagtgaatac ctttactggt gtgaagataa ttatgcacaa accctcacaa
2280

tacgcgttaa cattgaaacc tgtgaaatgt ccttaggttg ggtcatataa agccaaccat
2340

ttttgaggac catgtaccta gtgctttgaa aactgtaagt cactatatga atatgacaat
2400

atgtgcacat ttaaaattca gagctcggca ttgtgatact gatgcagaag ctagtagatt
2460

ggttaaaagt ctggacttct gtggcatttt tttcgtgacg tgataatcta tcataagcag
2520

acctaagcac agttttatga acacaatttt gcccatgaca ttgcctacag gatttccaga
2580

tgtgacttgc actcagaaga tcagtggtca acttcagaag ctcttccacg cttagatcat
2640

gtcttcagaa cttagatgtg aaaatctaca cactgggaga tgctgtgagc cccaaggttt
2700

tgatggagtt tgcttggaat cctcttgact tcatgccaca ttgacgtgaa ctttgatgta
2760

taataagcag cagcaacttc atgtgaaaat atggtcaggt agttatatgt aaggttacgt
2820

ggtccagtaa tgtcttagat tgataaatta ggtatggaat ccatcagtgt tacgtgatga
2880

gaataggtga acacaccttg tcagtgatga tgtaaacttc tctccttggc agggcatggg
2940

caaacatgct gattggtgca aatgtggtgc cgagctgtcc atagctgcag tgaaaggtga
3000

agagcaagac cttctctagg ttttctagct ttcattaaat gtatttttt ccccagagct
3060

aatttgaaag ttgattggac cactgtggat ggggtctcat taagaatgtg ggaaataggg
3120

gccgagtgcg gtggctcaca cctgtaatcc cagcagtttg gaaggccagg gcaggtggat
3180

cgcttgatcc caggaggtcg agaccagcct ggggaacaca tcctgtctct acaaaaaata
3240

caaaaattag ccaggcaggg tggtgcatgc ctgtagtccc agctacttgg gaggctgagg
3300

caggagaatt ttttgagccc aggatgcaga ggttgaagtg agccaagatc gtgccactgc
3360

actccagcct tgagacagag cgagaccctg tctcaaaaaa aaaaaaagaa cgtgggaaat
3420

atgaaccttt gaaagttaat ctgtgaattg aaagtttaac aataaaagta gttgtttgtt
3480

tcctttgg
3488

<210> 32

<211> 993

<212> DNA

<213> Homo sapiens

<400> 32
atggctgctg ttgacagttt ctacctcttg tacagggaaa tcgccaggtc ttgcaattgc
60

tatatggaag ctctagcttt ggttggagcc tggtatacgg ccagaaaaag catcactgtc
120

atctgtgact tttacagcct gatcaggctg cattttatcc cccgcctggg gagcagagca
180

gacttgatca agcagtatgg aagatgggcc gttgtcagcg gtgcaacaga tgggattgga
240

aaagcctacg ctgaagagtt agcaagccga ggtctcaata taatcctgat tagtcggaac
300

gaggagaagt tgcaggttgt tgctaaagac atagccgaca cgtacaaagt ggaaactgat
360

attatagttg cggacttcag cagcggtcgt gagatctacc ttccaattcg agaagccctg
420

aaggacaaag acgttggcat cttggtaaat aacgtgggtg tgttttatcc ctacccgcag
480

tatttcactc agctgtccga ggacaagctc tgggacatca taaatgtgaa cattgccgcc
540

gctagtttga tggtccatgt tgtgttaccg ggaatggtgg agagaaagaa aggtgccatc
600

gtcacgatct cttctggctc ctgctgcaaa cccactcctc agctggctgc attttctgct
660

tctaaggctt atttagacca cttcagcaga gccttgcaat atgaatatgc ctctaaagga
720

atctttgtac agagtctaat ccctttctat gtagccacca gcatgacagc acccagcaac
780

tttctgcaca ggtgctcgtg gttggtgcct tcgccaaaag tctatgcaca tcatgctgtt
840

tctactcttg ggatttccaa aaggaccaca ggatattggt cccattctat tcagtttctt
900

tttgcacagt atatgcctga atggctctgg gtgtggggag caaatattct caaccgttca
960

ctacgtaagg aagccttatg ctgcacagcc tga
993

<210> 33

<211> 330

<212> PRT

<213> Homo sapiens

<400> 33

Met Ala Ala Val Asp Ser Phe Tyr Leu Leu Tyr Arg Glu Ile Ala Arg
1 5 10 15

Ser Cys Asn Cys Tyr Met Glu Ala Leu Ala Leu Val Gly Ala Trp Tyr
20 25 30

Thr Ala Arg Lys Ser Ile Thr Val Ile Cys Asp Phe Tyr Ser Leu Ile
35 40 45

Arg Leu His Phe Ile Pro Arg Leu Gly Ser Arg Ala Asp Leu Ile Lys
50 55 60

Gln Tyr Gly Arg Trp Ala Val Val Ser Gly Ala Thr Asp Gly Ile Gly
65 70 75 80

Lys Ala Tyr Ala Glu Glu Leu Ala Ser Arg Gly Leu Asn Ile Ile Leu
85 90 95

Ile Ser Arg Asn Glu Glu Lys Leu Gln Val Val Ala Lys Asp Ile Ala
100 105 110

Asp Thr Tyr Lys Val Glu Thr Asp Ile Ile Val Ala Asp Phe Ser Ser

115 120 125

Gly Arg Glu Ile Tyr Leu Pro Ile Arg Glu Ala Leu Lys Asp Lys Asp
130 135 140

Val Gly Ile Leu Val Asn Asn Val Gly Val Phe Tyr Pro Tyr Pro Gln
145 150 155 160

Tyr Phe Thr Gln Leu Ser Glu Asp Lys Leu Trp Asp Ile Ile Asn Val
165 170 175

Asn Ile Ala Ala Ala Ser Leu Met Val His Val Val Leu Pro Gly Met
180 185 190

Val Glu Arg Lys Lys Gly Ala Ile Val Thr Ile Ser Ser Gly Ser Cys
195 200 205

Cys Lys Pro Thr Pro Gln Leu Ala Ala Phe Ser Ala Ser Lys Ala Tyr
210 215 220

Leu Asp His Phe Ser Arg Ala Leu Gln Tyr Glu Tyr Ala Ser Lys Gly
225 230 235 240

Ile Phe Val Gln Ser Leu Ile Pro Phe Tyr Val Ala Thr Ser Met Thr
245 250 255

Ala Pro Ser Asn Phe Leu His Arg Cys Ser Trp Leu Val Pro Ser Pro
260 265 270

Lys Val Tyr Ala His His Ala Val Ser Thr Leu Gly Ile Ser Lys Arg
275 280 285

Thr Thr Gly Tyr Trp Ser His Ser Ile Gln Phe Leu Phe Ala Gln Tyr
290 295 300

Met Pro Glu Trp Leu Trp Val Trp Gly Ala Asn Ile Leu Asn Arg Ser
305 310 315 320

Leu Arg Lys Glu Ala Leu Cys Cys Thr Ala
325 330

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 April 2005 (07.04.2005)

PCT

(10) International Publication Number
WO 2005/030985 A3

(51) International Patent Classification[7]: G01N 33/50, C12N 9/10

(21) International Application Number: PCT/GB2004/004103

(22) International Filing Date: 24 September 2004 (24.09.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/505,970 25 September 2003 (25.09.2003) US
0322493.8 25 September 2003 (25.09.2003) GB

(71) Applicant *(for all designated States except US)*: DEVGEN N.V. [BE/BE]; Technologiepark 30, B- 9052 Gent-Zwijnaarde (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: DE WILDE, Gert Jules Hector [BE/BE]; Dr. Armand Rubbensstraat 25, B- 9240 Zele (BE). SAUNDERS, Michael John Scott [GB/BE]; 132 rue Berkendael, B- 1050 Brussels (BE).

(74) Agents: BALDOCK, Sharon, Claire et al.; BOULT WADE TENNANT, Verulam Gardens, 70 Gray's Inn Road, LONDON WC1X 8BT (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

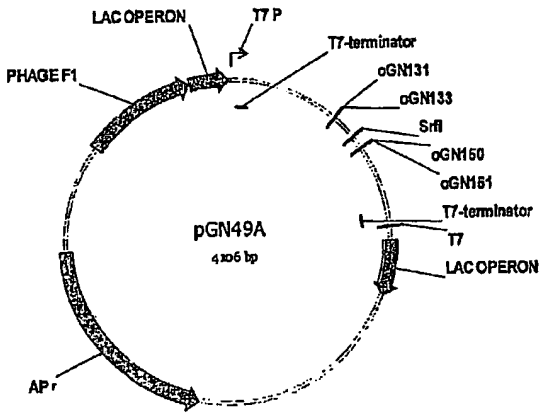
Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *with international search report*

(88) Date of publication of the international search report: 17 November 2005

[Continued on next page]

(54) Title: USE OF AMINO ACID SEQUENCES INVOLVED IN THE ELONGATION OF FATTY ACIDS IN IDENTIFYING AND/OR DEVELOPING COMPOUNDS FOR PREVENTING AND/OR TREATING METABOLIC DISEASES



(57) Abstract: The present invention relates to methods for the identification and/or the development of compounds that can be used to prevent and/or to treat metabolic diseases, and to the use of amino acid sequences that are involved in the elongation of fatty acids in such methods. The invention also relates to compounds that have been identified and/or developed using said methods, to pharmaceutical compositions that contain such compounds, and to the use of said compounds in the preparation of pharmaceutical compositions, in particular of pharmaceutical compositions for the prevention and/or treatment of metabolic diseases. The invention further relates to compounds that can interact with amino acid sequences that are involved in the elongation of fatty acids, to pharmaceutical compositions that contain such compounds, and to the use of said compounds in the preparation of pharmaceutical compositions, in particular of pharmaceutical compositions for the prevention and/or treatment of metabolic diseases.

WO 2005/030985 A3

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

Int[illegible]l Application No
PCT/GB2004/004103

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01N33/50 C12N9/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 02/062974 A (BAYER AKTIENGESELLSCHAFT; ZHU, ZHIMIN) 15 August 2002 (2002-08-15) abstract; sequence 2 claims 10-12,36,45 page 23, line 15 - page 24, line 6 page 36, line 19 - page 43, line 2 -----	1-12 13-19
X	WO 00/76308 A (EXELIXIS, INC; COSTA, MICHAEL, A; DOBERSTEIN, STEPHEN, KOHL; ELSON, SA) 21 December 2000 (2000-12-21) abstract claims 1,12,14,18-21 example 3 -----	1,2
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 February 2005	12-07-2005

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Vanhalst, K

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

Inter al Application No
PCT/GB2004/004103

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 03/019146 A (XENON GENETICS INC; WINTHER, MICHAEL, D; GRAY-KELLER, MARK, P) 6 March 2003 (2003-03-06) abstract claims 1-54	1,2
Y	MOON YOUNG-AH ET AL: "Identification of two mammalian reductases involved in the two-carbon fatty acyl elongation cascade." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 278, no. 9, 28 February 2003 (2003-02-28), pages 7335-7343, XP002316434 ISSN: 0021-9258 abstract RNAi-mediated inhibition of KAR and TER page 7339, column 2, paragraph 2 - column 1, paragraph 2	13-19
A	MOON Y-A ET AL: "Identification of a Mammalian Long Chain Fatty Acyl Elongase Regulated by Sterol Regulatory Element-binding Proteins" JOURNAL OF BIOLOGICAL CHEMISTRY 30 NOV 2001 UNITED STATES, vol. 276, no. 48, 30 November 2001 (2001-11-30), pages 45358-45366, XP002963525 ISSN: 0021-9258 abstract	1-19

INTERNATIONAL SEARCH REPORT

tional application No.
PCT/GB2004/004103

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-19 (partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-19 (partialy)

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 3 (human ELOVL6) or its mouse orthologue LCE.

---

2. claims: 1-19 (partially)

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 5 (further putative human enzyme involved in fatty acid elongation)

---

3. claims: 1-19 (partially)

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 8 (human ELOVL3) or its mouse orthologue Cig30.

---

4. claims: 1-19 (partially)

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 11 (human ELOVL5)

---

5. claims: 1-19 (partially)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 14 (human ELOVL4)

---

6. claims: 1-19 (partially)

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 17 (human ELOVL2) or its mouse orthologue Ssc2.

---

7. claims: 1-19 (partially)

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 20 (human ELOVL1) or its mouse orthologue Ssc1.

---

8. claims: 1-19 (partially)

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 23 (human KAR)

---

9. claims: 1-19 (partially)

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 26 (human TER)

---

10. claims: 1-19 (partially)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Method for identifying a compound useful for prevention or treatment of a metabolic disorder comprising : contacting a test chemical with an amino acid sequence involved in the elongation of faty acids, or a host cell or animal expressing such a sequence, and detecting interaction between both; wherein the amino acid sequence is that of SEQ ID n° 33 (homologue of KAR)

INTERNATIONAL SEARCH REPORT

formation on patent family members

Int[illegible]l Application No
PCT/GB2004/004103

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 02062974 A	15-08-2002	WO 02062974 A2	15-08-2002
WO 0076308 A	21-12-2000	US 6781028 B1	24-08-2004
		AU 5477000 A	02-01-2001
		CA 2373628 A1	21-12-2000
		EP 1196026 A1	17-04-2002
		JP 2003501102 T	14-01-2003
		WO 0076308 A1	21-12-2000
WO 03019146 A	06-03-2003	EP 1429784 A2	23-06-2004
		JP 2005500857 T	13-01-2005
		WO 03019146 A2	06-03-2003
		US 2003129129 A1	10-07-2003
		US 2004197847 A1	07-10-2004